# engineering design

## A PROJECT-BASED INTRODUCTION

4th edition

CLIVE L. DYM PATRICK LITTLE ELIZABETH J. ORWIN

WILEY

FOURTH EDITION

# *ENGINEERING DESIGN: A PROJECT-BASED INTRODUCTION*

**CLIVE L. DYM, PATRICK LITTLE, and ELIZABETH J. ORWIN**

*Harvey Mudd College*

WILEY

| | |
|---|---|
| VP & PUBLISHER | Don Fowley |
| EDITOR | Dan Sayre |
| EDITORIAL ASSISTANT | Jessica Knecht |
| MARKETING MANAGER | Chris Ruel |
| MARKETING ASSISTANT | Marissa Carroll |
| COVER DESIGNER | Miriam Dym |
| PHOTO EDITOR | Felicia Ruocco |
| ASSOCIATE PRODUCTION MANAGER | Joyce Poh |
| PRODUCTION EDITOR | Jolene Ling |

This book was set by Thomson Digital. Cover and text printed and bound by Edwards Brothers Malloy.

This book is printed on acid free paper.

Founded in 1807, John Wiley & Sons, Inc. has been a valued source of knowledge and understanding for more than 200 years, helping people around the world meet their needs and fulfill their aspirations. Our company is built on a foundation of principles that include responsibility to the communities we serve and where we live and work. In 2008, we launched a Corporate Citizenship Initiative, a global effort to address the environmental, social, economic, and ethical challenges we face in our business. Among the issues we are addressing are carbon impact, paper specifications and procurement, ethical conduct within our business and among our vendors, and community and charitable support. For more information, please visit our website: www.wiley.com/go/citizenship.



Library of Congress Cataloging-in-Publication Data

Dym, Clive L.
Engineering design : a project-based introduction / Clive L. Dym, Patrick Little and Elizabeth J. Orwin, Harvey Mudd College. – 4th edition.
pages cm
Includes bibliographical references and index.
ISBN 978-1-118-32458-5 (pbk.)
1. Engineering design. I. Little, Patrick, 1952- II. Orwin, Elizabeth J. III. Title.
TA174.D958 2014
620'.0042–dc23

2013016211

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

To
Joan Dym
*whose love and support are distinctly nonquantifiable*

cld

Charlie Hatch
*a teacher's teacher*

pl

Carl Baumgaertner
*who inspired me to teach*

ejo

# CONTENTS

# *FOREWORD**

To design is to imagine and specify things that don't exist, usually with the aim of bringing them into the world. The "things" may be tangible—machines and buildings and bridges; they may be procedures—the plans for a marketing scheme or an organization or a manufacturing process, or for solving a scientific research problem by experiment; they may be works of art—paintings or music or sculpture. Virtually every professional activity has a large component of design, although usually combined with the tasks of bringing the designed things into the real world.

Design has been regarded as an art, rather than a science. A science proceeds by laws, which can sometimes even be written in mathematical form. It tells you how things must be, what constraints they must satisfy. An art proceeds by heuristic, rules of thumb, and "intuition" to search for new things that meet certain goals, and at the same time meet the constraints of reality, the laws of the relevant underlying sciences. No gravity shields; no perpetual motion machines.

For many years after World War II, science was steadily replacing design in the engineering college curricula, for we knew how to teach science in an academically respectable, that is, rigorous and formal, way. We did not think we knew how to teach an art. Consequently, the drawing board disappeared from the engineering laboratory—if, indeed, a laboratory remained. Now we have the beginnings—more than the beginnings, a solid core—of a science of design.

One of the great gifts of the modern computer has been to illuminate for us the nature of design, to strip away the mystery from heuristics and intuition. The computer is a machine that is capable of doing design work, but in order to learn how to use it for design, an undertaking still under way, we have to understand what the design process is.

We know a good deal, in a quite systematic way, about the rules of thumb that enable very selective searches through enormous spaces. We know that "intuition" is our old friend "recognition," enabled by training and experience through which we acquire a great collection of familiar patterns that can be recognized when they appear in our problem situations. Once recognized, these patterns lead us to the knowledge stored in our memories. With this understanding of the design process in hand, we have been able to reintroduce design into the curriculum in a way that satisfies our need for rigor, for understanding what we are doing and why.

One of the authors of this book is among the leaders in creating this science of design and showing both how it can be taught to students of engineering and how it can be implemented in computers that can share with human designers the tasks of carrying out the design process. The other is leading the charge to integrate the management sciences into both engineering education and the successful conduct of engineering design projects. This book thus represents a marriage of the sciences of design and of management. The science of design continues to move rapidly forward, deepening our understanding and enlarging our opportunities for human-machine collaboration. The study of design has joined the study of the other sciences as one of the exciting intellectual adventures of the present and coming decades.

*Herbert A. Simon*
Carnegie Mellon University
Pittsburgh, Pennsylvania
August 6, 1998

* Herb Simon graciously contributed the foreword for our first edition. Unfortunately, the passage of time since was marked by the loss of one of our great heroes and a true renaissance mind: Herb passed away on March 4, 2002. We still feel the loss.

# PREFACE

When we started on the first edition of this book in the late 1990s, we could not have predicted that we would someday be asked to prepare a fourth edition of a text for a then-controversial course. At that time, a cornerstone introduction to engineering design was indeed considered improbable, if not impossible or meaningless. Now such courses are a staple of many engineering programs, and we are proud to have helped bring that curricular adaptation to life. We have also been part of a similar adaptation of engineering's capstone courses, which were then often undertaken more in response to accreditation needs than a desire for real-world projects. Today externally focused capstone courses, some modeled on Harvey Mudd College's Engineering Clinic, not only give students an authentic design experience, but also often introduce them to working with peers scattered around the world. The students in the classroom or design studio have also changed: Many more women and underrepresented minority students now major in engineering.

These transitions have been accompanied by an evolution in the discipline of design and in the perception of engineering design by the faculties of engineering schools. In particular, design is now a recognized intellectual discipline, with a vocabulary, structure, and methods that reflect our increasing ability to articulate what we are doing when we design something. And as with many other disciplines, design ranges from the narrow and mathematical (e.g., kinematics, optimization) to the broad and transdisciplinary (e.g., the life of a product from its inception to use to disposal, the communication and teamwork skills that are the "soft" skills of engineering design).

We have also changed, certainly getting older, perhaps also becoming wiser. We have had opportunities to see how the design ideas we taught worked, which needed refinement, and which didn't work at all. We have tried to adapt this fourth edition both to the changing circumstances and to our increased knowledge of the world, the engineering profession, and our educational mission.

Of course, some things have not changed at all. Engineering design has always required attention to the wishes of the client, users, and the larger public. It is still true that engineers must organize their design processes to communicate their design thinking to their design partners. And it also remains true that effective design teams are those whose members respect one another. Perhaps most of all, a commitment to ethical design by and on behalf of a diverse community must remain at the forefront of what it is we do as engineers.

Today there are many more books on design, engineering design, project management, team dynamics, project-based learning, and the other topics we cover in this volume, than when we wrote our first edition. We wanted then—as we still do today—to combine these topics in a single, introductory work that focused particularly on conceptual design. That original desire arose from our teaching at Harvey Mudd College, where our students do team-based design projects in a first-year design course, *E4: Introduction to Engineering Design* (called "E4"), and in the Engineering Clinic. Clinic is an unusual capstone course taken by juniors (for one semester) and seniors (for both semesters) in which students work on externally sponsored design and development projects. In both E4 and Clinic, Mudd students work in multidisciplinary teams, under specified time deadlines, and within specified budget constraints. These conditions are meant to replicate to a significant degree the environments within which most practicing engineers will do much of their professional design work. In looking for books that could serve our audience, we found that there were excellent texts covering detailed design, usually targeted toward senior capstone design courses, or "introductions to engineering" that focused on describing the branches of engineering. We could not find a book that introduced the processes and tools of conceptual design in a project or team setting that we found suitable for first- and second-year students. And while other more "skills-oriented" texts and series have come onto the market since, we are gratified that a growing market has emerged for the book that addresses our original concerns.

In designing all four editions of this book, we confronted many of the same issues that we discuss in the pages that follow. It was important for us to be very clear about our overall objectives, which we outline below, and about the particular objectives we had for each chapter. We asked about the pedagogic function served by the various examples, and whether some other example or tool might provide a better means for achieving that pedagogical function. The resulting organization and writing represent our implementation of our best design. Thus, this and all books are designed artifacts: They require the same concern with objectives, choices, constraints, functions, means, budget, and schedule, as do other engineering or design projects.

This book is directed to three audiences: students, teachers, and practitioners. The book is intended to support *students* to learn about design, the central activity of engineering, by *doing* design. We view our design course, E4, as a setting in which students *acquire design skills* as they *experience the activity of design* by working on design projects. The book is intended to help students learn formal design tools and techniques as they solve conceptual design problems. They can then apply these formal methods to other design projects they will face later in their education in Clinic-like capstone courses and later in their careers. Students will also learn about communication, team dynamics, and project management. We have included examples of work done by our students on actual projects in E4, both to show how the tools are used and to highlight some frequently made mistakes.

We wrote this book with *teachers* also very much in mind. We thought about how to deliver the material to students, and about how introductory design courses could be taught. In this fourth edition, we decomposed and modularized much of the text, in order to avoid the confusion that often results when a new vocabulary is being learned; that is, to separate objectives from constraints, objectives from functions, functions from means, and customer requirements from design specifications. The modularization also provides options for instructors to structure their classes in a variety of ways, bringing forward (or deferring) discussions of communication, team dynamics, leadership, or management, because the chapters on these (and other) topics are self-contained. We also provide a complete design case study and two continuing design examples that can be used by an instructor as ongoing examples for illustration and as in-class exercises. (We don't assign homework problems in E4 as our students are working on their various E4 projects as "homework" when they're not in class.) In an accompanying *Instructor's Manual*, we outline sample syllabi and organizations for teaching the material in the book, as well as additional examples.

Finally, we hope the book will be useful to *practitioners*, either as a refresher of things learned or as an introduction to some essential elements of conceptual design that were not formally introduced in engineering curricula in years past. We do not assume that the case study or the illustrative design examples given here substitute for an engineer's experience, but we do believe that they show the relevance of these tools to practical engineering settings. Some of our friends and colleagues in the profession like to point out that the tools we teach would be unnecessary if only we all had more common sense. Notwithstanding that, the number and scale of failed projects suggest that common sense may not, after all, be so commonly distributed. In any case, this book offers both practicing engineers (and engineering managers) a view of the design tools that even the greenest of engineers will have in their toolbox in the coming years.

## SOME REMARKS ON VOCABULARY AND WORD USAGE

There is no engineering design community that transcends all engineering disciplines or all types of engineering practice. For that very reason, words are used differently in different domains, and so differing technical jargons have developed. Since we want to provide a unified coherent understanding that would be a useful foundation for all of our students' future design work, whether in their formal studies or in their chosen careers, we begin our discussions of the major concepts and terms of art with formal dictionary definitions, but leavened by our understanding of today's "best practices" in design. We do this to remind readers that word usage has its roots in a shared understanding of vocabulary, in our case the English vocabulary. Even technical jargon has—or should have—a traceable path back to common usage. Thus, in this fourth edition we have worked much harder than we have before to be as crisp and consistent as possible with the words we chose to use.

Further, it is clear that words are used differently in the different domains of engineering practice. For example, different authors (in both the research literature and textbooks) define phases of the design process differently, with varying activities occurring within them. We have worked very hard to clearly articulate our model of the design process in Chapter 2. As we reviewed materials for this edition, we saw that the use of the terms *requirements* and *specifications* in engineering practice is not uniform. Thus, we choose to speak in terms of *customer requirements* to specify what the client wants and needs from her design (i.e., the client's *objectives* and *constraints* and the *functions* as she'd like them to happen), and *design specifications* to articulate in engineering terms how a design is supposed to perform its *functions* and, as appropriate, display its *behaviors*.

## SOME SPECIFICS ABOUT WHAT'S COVERED

Design is an *open-ended* and *ill-structured* process, by which we mean there is no unique solution, and that the candidate solutions cannot be generated with an algorithm. As we emphasize in the early chapters, designers have to provide an orderly process for organizing an ill-structured design activity in order to support making decisions and trade-offs among possibly competing solutions. In such cases, algorithms and mathematical formulations cannot replace the imperative to understand the often subjective needs of various stakeholders (clients, users, the public, and so on)—even if those mathematical tools are used later in the design process. Perhaps ironically, this lack of structure and the inapplicability of formal mathematical tools make the introduction of conceptual design early in the curriculum possible and, we think, desirable. It provides a framework in which engineering science and analysis can be used, while not demanding skills that most first- and second-year students have not yet acquired. We have, therefore, included in this book the following specific tools for conceptual design, for acquiring and organizing design knowledge, and for managing the team environment in which design takes place.

The following *formal conceptual design methods* are delineated:

- objectives trees
- establishment of metrics to measure the achievement of objectives
- pairwise comparison charts (PCCs) to rank objectives
- functional analysis (including black and glass boxes, enumeration, function-means trees, and so on)
- morphological ("morph") charts to develop design alternatives
- specifications development

Since both the framing or defining of a design problem and conceptual design thinking require and produce a lot of information, we introduce a variety of means to acquire and process information, including literature reviews, brainstorming, analogies, user surveys and questionnaires, reverse engineering (or dissection), simulation and computer analysis, and formal design reviews.

The successful completion of any design project by a team requires that team members estimate a project's scope of work, schedule, and resources early in the life of the project. To this end, we introduce several *design management tools*:

- work breakdown structures (WBSs)
- schedules
- budgets

We also discuss several other topics that we feel are increasingly important in a first exposure to design. We discuss the completion of a design project, with a strong emphasis on the *ways and means of reporting design results* in Chapters 9 and 10. These chapters allow instructors to focus on engineering communication as an integral part of the design process, including engineering drawings, reports, and presentations. We also present some more practical aspects of drawing and tolerancing in Appendix A. We did this because we wanted to bring together the basic skills needed in design, such as communicating through drawings by adhering to appropriate standards and conventions (e.g., geometric dimensioning and tolerances).

We also include a discussion about *building physical models and prototypes* in Chapter 11. We did this because we have also observed in our own students that most don't start college with much hands-on experience, even in basic woodcraft. Since we expect them to build elementary (physical) models and prototypes, it seemed only fair to include some understanding of what models and prototypes are, as well as (in Appendix B) some cautionary tips about working in a shop or laboratory, and some very basic tips on how to actually make (and fasten) some basic wooden parts.

In Chapter 12, we introduce some ideas about *mathematical modeling in design*, placed in the context of doing preliminary and detailed design. The material introduces principles of mathematical modeling to reinforce concepts behind applying mathematics and physics to engineering. Then we go on to illustrate a few of the kinds of calculations that might be done in the later phases of design. We illustrate the modeling of both battery-powered payload carts and a basic rung or step for a ladder, where we apply some results from elementary beam theory. Needless to say, in one chapter and in the kinds of course that we aimed this book toward, we could not delve into preliminary and detailed design in all engineering disciplines. What we present is representative of the "good habits of thought" needed to model and analyze designs in all disciplines.

In Chapter 13 we present a brief introduction to engineering economics and to the time value of money, the latter being quite important because we often need to balance initial or present costs against costs due, for example, to use, wear, and maintenance. In Chapter 14 we discuss "design for *X*" issues, including use, manufacturing and assembly, reliability and maintainability, and sustainability. This chapter provides a vehicle for faculty who want to expand on these topics and lead students into issues such as concurrent design, DFM, or emerging areas such as sustainability and carbon footprints.

In Chapter 15 we undertake a discussion of teams, exploring both the stage of team formation and the roles of individuals on both effective and ineffective teams. Then in Chapter 16 we talk about the fundamentals of managing a design project, including monitoring its progress and controlling its expenditures and costs. We finish our exploration of engineering design with our own capstone, Chapter 17, in which we discuss important ethics issues in design. This chapter reflects a wider notion of engineering ethics than in the past, as we invite faculty to address traditional notions of liability and responsibility and also newer ideas of social and political dimensions of engineering design.

## DESIGN CASE STUDY AND INTEGRATIVE DESIGN EXAMPLES

We use one case study and two integrative examples to follow the design process through to completion, thus showing each of the tools and techniques as they are used on a design project. In addition to numerous "one-time" examples, we detail the following case study and integrative examples:

***Design case study***: This case study, contained in full in Chapter 2, follows the design of a ***microlaryngeal surgical stabilizer***, a device used to stabilize the physician's hand as he uses various instruments in throat surgery. The work we show in this case study derives from the efforts of several student teams in the Harvey Mudd College's first-year design course ("E4"), on a project sponsored by the Beckman Laser Institute of the University of California at Irvine. (Further details can be found in the Acknowledgments, the Notes at the end of Chapter 2 and the References and Bibliography.)

The first **illustrative design example** is the design of a ***juice container***. This is a design project created by the authors solely to illustrate the application of various conceptual design tools that are the substance of much of this book. A design team, having a fruit juice company as a client, is asked to develop a means of delivering a new juice to a market predominantly composed of children and their parents. There are clearly a number of possibilities (e.g., mylar bags, molded plastics), and issues such as environmental effects, safety, and the costs of manufacturing are considered.

The second **illustrative design example 2** is the design of an ***arm support*** to be used by a child diagnosed with cerebral palsy (CP). Here we show how teams of Harvey Mudd College students in our E4 design class responded to the challenge of designing something for one such disabled student, having in mind at the same time that such a design might be useful to many other children in many other schools. We show work done by two particular teams, again to illustrate how these student teams applied the design tools they were learning. (Again, further details can be found in the Acknowledgments, the Notes at the end of Chapter 2, and the References and Bibliography.) Prototypes were subsequently built by the students and delivered to the Danbury School, a special education elementary school within the Claremont Unified School District of Claremont, California.

Finally, an accompanying *Instructor's Manual* includes a case study of the design of a ***transportation network*** to enable automobile commuter traffic between Boston and its northern suburbs, through Charlestown, Massachusetts. This conceptual design problem clearly illustrates the many factors that go into large-scale engineering projects in their early stages, when choices are being made between highways, tunnels, and bridges. Among the design concerns are cost, implications for future expansion, and preservation of the character, environment, and even the view of the affected neighborhoods. This project is also an example of how conceptual design thinking can significantly influence some very "real-world" events.

As noted at the outset, this edition has presented both an opportunity and a challenge for us as authors. We now share those with our readers.

*Clive L. Dym*
*Patrick Little*
*Elizabeth J. Orwin*
Claremont, California
March 7, 2013

# ACKNOWLEDGMENTS

A book like this does not get written without the support, advice, criticism, and help of many people. We continue to be grateful to the many colleagues and friends who helped us bring the first three editions to fruition, and our thanks were detailed in those prior editions in 1997, 2004, and 2009.

For this fourth edition, we want to add the following thanks to:

The *HMC E4 student design teams* that developed the design work products that we used as a case study and as illustrative design examples. Those teams and their projects are listed in our bibliography as (Ahmad et al. 2007), (Attarian et al. 2007), (Best et al. 2007), (Both et al. 2000), (Chan et al. 2000), (Feagan et al. 2000), and (Saravanos et al. 2000).

*Miriam Dym* for designing the cover for this fourth edition.

*Dan Sayre* of John Wiley & Sons, our editor and steadfast champion, and *Jessica Knecht*, his very able, very helpful, and always gracious assistant.

*R. Erik Spjut*, our colleague at HMC, for providing several drawings and some materials on preliminary design and insights into practical building.

*Bob Welsh*, Vice President of Engineering of DeWalt Tools, for providing the exploded graphic of the DeWalt DW21008K corded power drill.

*Michael A. Sandford*, President of Technical Documentation Center of Arizona, for permission to reprint several figures in Appendix B.

The *American Society of Civil Engineers*, for permission to reprint portions of ASCE's Code of Ethics.

The *American Society of Mechanical Engineers*, for permission to reprint portions of ASME's Code of Ethics and several figures from ASME Y14.3-1975 and ASME Y14.4M-1994 (R2004).

The late *Herbert A. Simon* of Carnegie Mellon University, whose foreword (written for the very first edition and reprinted in each subsequent edition) and generous encouragement of CLD still provide inspiration.

*Amos G. Winter*, *Daniel D. Frey*, and *Global Research Innovation and Technology* (GRIT) of the Massachusetts Institute of Technology, for providing a photo of the Freedom Wheelchair.

Finally, *to each and all of our spouses and families*, for tolerating us during the absences that such projects entail, and for listening to each of us as we worked through differences to find a common voice.

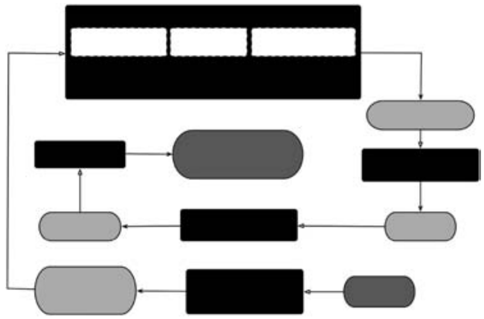
# PART I

# *INTRODUCTION*

CHAPTER 1

# ENGINEERING DESIGN

*What does it mean to design something? Is engineering design different from other kinds of design?*



PEOPLE HAVE been designing things for as long as we can archaeologically uncover. Our earliest ancestors designed flint knives and other tools to help meet their most basic needs. Their wall paintings were designed to tell stories and to make their primitive caves more attractive. Given the long history of people designing things, it is useful to set some context for engineering design and to start developing a vocabulary and a shared understanding of what we mean by engineering design.

## 1.1 WHERE AND WHEN DO ENGINEERS DESIGN?

What does it mean for an *engineer* to design something? When do engineers design things? Where? Why? For whom?

An engineer working for a large company that processes and distributes various food products could be asked to design a container for a new juice product. She could work for a design-and-construction company, designing part of a highway bridge embedded in a larger transportation project, or for an automobile company that is developing new instrumentation clusters for its cars, or for a school system that wants to design specialized facilities to better serve students with orthopedic disabilities.

There are common features that make it possible to identify a design process and the context in which it occurs. In each of these cases, three "roles" are played as the design unfolds. First there is a *client*, a person or group or company that wants a design conceived. There is also a *user* who will employ or operate whatever is being designed. Finally, there is a *designer* whose job is to solve the client's problem in a way that meets the user's needs. The client could be internal (e.g., a person at the food company in charge of the new juice product) or external (e.g., the government agency that contracts for the new highway system). While a designer may relate differently to internal and external clients, it is typically the client who *motivates* and presents the starting point for design. That is why a designer's first task is to *question* the client to clarify what the client really wants and translate it into a form that is useful to her as an engineer. We'll say more about this in Chapter 3 and beyond.

It is worth noting that the client, the user, and even the designer may not always be three or even two different people: In a small start-up, for example, the designer may be the client, and may also rely on his or her own personal experience as a user when initiating a design. Similarly, for an internal project, the roles may again merge. However, for most design projects, it is useful to distinguish between the three roles and their respective responsibilities—as anyone who has used beta versions of software can testify because all too often, software designers imagine that their own experience is sufficient for every user!

The user is a key player in the design effort. In the contexts mentioned above, the users are, respectively, consumers who buy and drink a new juice drink, drivers on a new interstate highway, and students with orthopedic disabilities (and their teachers). Users have a stake in the design process because designs have to meet their needs. Thus, the designer, the client, and the user form a triangle, as shown in Figure 1.1. The designer has to understand what both the client and users want and need. Often the client speaks to the designer on behalf of the intended users, although anyone who has sat in a cramped seat on a commercial flight would have to ask both airlines and airplane manufacturers who they think their users are!

The *public* also has a stake in many designs, for example, a new interstate highway. While the notion of the public may seem to be implicit in the user, this is not always the case. Explicitly identifying who is affected by a design is important, because it may raise ethical issues in design projects, as we will explore in Chapter 17.

It is clear that both designer and client *have to understand what the users want and what the public demands in a design*. In Chapter 2, we will describe design processes that model how engineers interact with and communicate their design thinking to clients and potential users. In Chapters 3–5, we will identify some tools to organize and refine that thinking.

Engineering designers work in many different kinds of environments: small and large companies, start-up ventures, government, not-for-profit organizations, and engineering

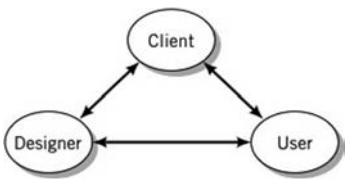



**Figure 1.1** The designer–client–user triangle shows three parties involved in a design effort: a client, who has objectives that must be realized; the users of the design, who have their own wishes; the designer, who must design something that can be built and that satisfies everybody.

services firms. Designers will see differences in the size of a project, the number of colleagues on the design team, and their access to relevant information about what users want. On large projects, many designers will be working on details of a project that are so confined that much of what we describe in this book may not seem immediately useful. The designers of a bridge abutment, an airplane fuel tank, or components of a computer motherboard are not likely to be as concerned with the larger picture of what clients and users want from the entire project because the system-level design context has already been established. These are *detailed design* problems in which more general design issues have already been decided. However, all projects begin with *conceptual design*. Thinking about the size and mission of an airplane will have been done before fuel tank design begins, and the overall performance parameters of the computer motherboard will be determined prior to selecting specific chips.

Large, complex projects often lead to very different interpretations of client project statements and of user needs. One has only to look at the many different kinds of skyscrapers that decorate our major cities to see how architects and structural engineers envisage different ways of housing people in offices and apartments. Visible differences also emerge in airplane design (Figure 1.2) and wheelchair design (Figure 1.3). Each of these sets of devices could result from a simple, common design statement: Airplanes are


U.S. Army Corps of Engineers Sacramento District

Courtesy of USFWS

Matt York / Associated Press

© GA161076 / iStockphoto

**Figure 1.2** Several aircraft, each of which "safely transports people or goods through the air," and each of which was designed for a different mission.


Image Source / Getty Images

© 26ISO / iStockphoto

© Daniel Korzeniewski / Shutterstock.com

Amos Winter, Daniel Frey, and Global Research Innovation and Technology (GRIT)

**Figure 1.3** A collection of "personal mobility devices to transport people unable to use their legs," that is, a set of very different wheelchairs.

"devices to transport people and goods through the air," and wheelchairs are "personal mobility devices for people who are unable to use their legs." However, the different products that have emerged represent different concepts of what clients and users wanted (and what designers perceived they wanted) from these devices. Designers have to clarify what clients want and then translate those wants into an engineered product.

The designer–client–user triangle also prompts us to recognize that the interests of the three players might diverge and consider the consequences of such divergence. The presence of multiple interests creates an interaction of multiple obligations, and these obligations may conflict. For example, the designer of a juice container might consider metal cans, but easily "squashed" cans are a hazard if sharp edges emerge during the squashing. There could be trade-offs among design variables, including the material of which a container is to be made and the container's thickness. The choices made in the final design could reflect different assessments of the possible safety hazards, which in turn could lay a foundation for potential ethics problems. Ethics problems, which we will discuss in Chapter 17, occur because *designers have obligations not only to clients and users, but also to their profession and to the public at large*, as detailed in the codes of ethics of engineering societies. Thus, ethics issues are always part of the design process.

Another aspect of engineering design practice that is increasingly common in projects and firms of all sizes is that *teams* do design. Many engineering problems are inherently multidisciplinary (e.g., the design of medical instrumentation), so there is a need to understand the requirements of clients, users, and technologies in very different ways. This requires that teams be assembled to understand and address such different needs. The widespread use of teams clearly affects how design projects are managed, another recurring theme of this book.

Engineering design is a multifaceted subject. In this book, we offer a framework to facilitate productive thought about the conceptual issues and the resulting choices made early in the design of many different engineered products.

## 1.2 A BASIC VOCABULARY FOR ENGINEERING DESIGN

There are many definitions of *engineering design* in the literature, and there is a lot of variation in how engineers describe design actions and attributes. We will now define what *we* mean by engineering design and also some of the related terms that are commonly used by engineers and designers.

### 1.2.1 Defining Engineering Design

The following formal definition of engineering design is the most useful one for our purposes:

- **Engineering design** is a systematic, intelligent process in which engineers generate, evaluate, and specify solutions for devices, systems, or processes whose form(s) and function(s) achieve clients' objectives and users' needs while satisfying a specified set of constraints. In other words, *engineering design is a thoughtful process for generating plans or schemes for devices, systems, or processes that attain given objectives while adhering to specified constraints*.

It is important to recognize that when we are designing devices, systems, and processes, we are designing *artifacts*: artificial, manmade objects, the "things" or devices that are being designed. They are most often physical objects such as airplanes, wheelchairs, ladders, cell phones, and carburetors. But "paper" products (or their electronic versions) such as drawings, plans, computer software, articles, and books are also artifacts in this sense. In this text we will use device, artifact, or system rather interchangeably as the objects of our design.

With further recourse to our "design dictionary," we note the following definitions:

- **design objective** *n*: a feature or behavior that we wish the design to have or exhibit.
- **design constraint** *n*: a limit or restriction on the features or behaviors of the design. A proposed design is unacceptable if these limits are violated.

- **functions** *n*: things a designed device or system is supposed to do. Engineering functions almost always involve transforming or transferring energy, information, or material. We view energy transformation or transfer quite broadly: It includes supporting and transmitting forces, the flow of current, the flow of charge, the transfer of material, and so on.
- **means** *n*: a way or a method to make a function happen. For example, *friction* is a ***means*** of fulfilling a function of *applying a braking force*.
- **form** *n*: the shape and structure of something as distinguished from its material. We will not deal with form very much in this book, but form is central to industrial design, a very important part of product design.

Note that *objectives for a design are different from the constraints placed on a design*. Objectives may be completely or partially achieved, or may not be achieved at all. Constraints, on the other hand, *must be* satisfied or the design is not acceptable. That is, they are binary (yes or no): There are no intermediate states. If we were designing a corn degrainer for Nicaraguan farmers to be cheaply built of indigenous (local) materials, one objective might be to make it as cheap as possible, while a constraint might limit the cost to less than US$20.00. Making the degrainer of indigenous materials could be an objective if it is a *desired* attribute, or a constraint if it is a *required* attribute.

Our definition of engineering design states that designs emerge from a *systematic, intelligent process*. This is not to deny that design is a creative process. There are, however, techniques and tools we can use to support our creativity, to help us think more clearly, and to make better decisions along the way. These tools and techniques, which form much of this book, are not formulas or algorithms. Rather, they are ways of asking questions and of presenting and reviewing the answers to those questions as the design process unfolds.

### 1.2.2 Assumptions Underlying Our Definition of Engineering Design

There are some implicit assumptions behind our definition of engineering design and the terms in which it is expressed. It is useful to make them explicit.

First, *design is a thoughtful process that can be understood*, and therefore both taught and learned. Without meaning to spoil the magic of creativity or the importance of innovation in design, people *think* while designing. So it is important to have tools to support that thinking, to support design decision making and even design project management.

The *formal methods* we use to generate design alternatives follow naturally from our inclination to think about design. This might seem pretty obvious: There's not much point in considering new ways of looking at design problems or talking about them—unless we can exploit them to do design more effectively. Thus, our formal methods are part of the (formal) process we use to identify and clarify what a client *wants* (i.e., objectives), *needs* (i.e., constraints), and intends the *design to do* (i.e., its functions). We will describe such a process in Chapter 2, and we will show how it begins with a client's problem statement and ends with a *functionally complete* design that does everything the client wants it to do, has the desired attributes, and stays within the client's constraints.

### 1.2.3 Measuring the Success of an Engineered Design

How do we know whether our design is successful? We make measurements. What do we measure? Early in the design process we establish a set of *metrics* to ascertain or measure the extent to which a proposed design meets our design objectives:

- **metric** *n*: a standard of measurement; in the context of engineering design, a scale on which the achievement of a design's objectives can be measured and assessed.

*Metrics* provide scales or rulers on which we can measure the degree to which objectives are achieved. To offer a truly simple example, let us suppose an objective of being able to jump as far as possible. A metric for such a jump might be based on using a ruler to measure the distance jumped (in feet or meters). There are interesting issues that must be addressed when talking about metrics: All objectives are not easily quantified, their quantifications are not readily compared, and not all measurements are easily made. We discuss these issues in Chapter 4. *We will use metrics to mean rulers or standards specifically for objectives*.

Later in the design process, we establish *specifications* to express in engineering terms a design's functional behavior. Setting out such specifications is an essential aspect of the "best practices" of engineering design as it is currently done in industry:

- **specification(s)** *n*: a scale on which the achievement of a design's functions can be measured. Specifications are engineering statements of the extent to which functions are performed by a design.

Design specifications are stated in a number of different ways, depending on what the designer intends to articulate. Thus, specifications may specify *values* for particular functions or design features, *procedures* for calculating functions or behaviors of the design, or *performance levels* that must be attained by the design.

It is important to note that the vocabulary of design practice varies across different engineering disciplines and related fields such as computer science. In fact, the terms *specifications* and *requirements* are often taken as synonymous descriptors of a design's features and behaviors, as well as its functions. For the sake of clarity, we will, in Chapters 2 and 5, take a specific stance about these two terms, as follows: We will normally use *requirements* as shorthand for *customer requirements*, which are the client's statement of objectives, constraints, and functions. We will use *specifications* as shorthand for *engineering specifications* or *design specifications*, which are the designer's expression of what a design is intended to do in engineering terms. We will define requirements and specifications in greater detail in Chapter 2, and will explore the nature of design specifications extensively in Chapter 5.

### 1.2.4 Form and Function

*Form* and *function* are two related yet independent entities. This is important. We often think of the design process as beginning when we sit down to draw or sketch something, which suggests that form is a typical starting point. However, function is an altogether different aspect of a design that may not have an obvious relationship to its shape or form. In particular, while we can often infer the purpose of a device from its form or structure, we can't do the reverse, that is, we cannot automatically deduce what form a device must have *from the function alone*. To take a simple example, we can't look at the shape of a smartphone and know what it was supposed to do. Moreover, if we were asked to design a smartphone, is there any obvious link or inference that we can use to choose its form or shape? That is, knowing that we want to achieve the *function* of wireless telephony does not lead us to (or even suggest) any of the *forms* of smartphones.

### 1.2.5 Design and Systems

While our focus is on the design of "a thing," there are two broader issues that are worth thinking about, both having to do with *systems*. First, no thing or device stands alone, entirely independent of its environment: It usually works in some environment and often has to interface with other devices. Thus, a definition offered by the late Herbert A. Simon, Nobel laureate in economics and founding father to several fields, including design theory: "*Design* is an activity that intends to produce a description of an artifice in terms of its organization and functioning—its interface between inner and outer environments." This definition places designed objects in a *systems* context that recognizes that any artifact operates as part of a system that includes the world around it. In this sense, all design is systems design because devices, systems, and processes must each operate within and interact with their surrounding environments.

This leads to the second thought about design and systems. The major design challenges facing engineers in the decades to come will be less about devices of "stand-alone" artifacts, and more concerned with designing complex engineering systems. These have been defined as "a class of systems characterized by a high degree of technical complexity, social intricacy, and elaborate processes, aimed at fulfilling important functions in society." Examples of such complex systems include the U.S. interstate highway system, the country's electric power grid, and the Internet. Clearly there are many more issues involved (and things to be learned) in designing such large technical systems, but the problem definition and problem-solving approaches we introduce here will be useful in attacking them.

### 1.2.6 Communication and Design

Finally, our definition of engineering design and the related assumptions we have identified rely heavily on the central role of communication in the design process. Some set of languages or representations is involved in every part of the design process. From the original communication of a design problem, through the final fabrication specifications, the device or system being designed must be described and "talked about" in many, many ways. *Communication is a key issue*. It is not that problem solving and evaluation are less important; they are extremely important. But problem solving and evaluation are done at levels and in styles—whether spoken or written languages, numbers, equations, rules, charts, or pictures—that are appropriate to the immediate task at hand. Successful work in design is inextricably bound up with the ability to communicate.

Engineering designers do not typically produce their artifacts, except in the form of prototypes and proofs of concept. While these prototypes are useful for understanding the design space and demonstrating the feasibility of the design, the ultimate product of most contemporary design is a set of fabrication specifications for others to use in making the artifacts. These fabrication specifications provide a detailed description of the designed device so that it can be assembled or manufactured, thus separating the "designing" from the "making." This description must be both complete and quite specific; there should be no ambiguity and nothing can be left out. Indeed, this specification may be the only connection between a designer and the fabricator or maker of the design.

Traditionally, fabrication specifications were presented in a combination of drawings (e.g., detailed engineering drawings, circuit diagrams, flow charts) and text (e.g., parts lists, materials specifications, assembly instructions). We can achieve completeness and specificity with such traditional specifications, but we may not capture the designer's intent—and this can lead to catastrophe. In 1981, a suspended walkway across the central atrium in the Hyatt Regency Hotel in Kansas City collapsed because a contractor fabricated the connections for the walkways in a manner different than intended by the original designer.

In that design, walkways at the second and fourth floors were hung from the same set of threaded rods that would carry their weights and loads to a roof truss (see Figure 1.4). The fabricator was unable to procure threaded rods sufficiently long (i.e., 24 ft.) to suspend the second-floor walkway from the roof truss, so instead, he hung it from the fourth-floor walkway with shorter rods. (It also would have been hard to screw on the bolts over such lengths and attach walkway support beams.) The fabricator's redesign was akin to requiring that the lower of two people hanging independently from the same rope change his position so that he was grasping the feet of the person above. That upper person would then be carrying both people's weights with respect to the rope. In the hotel, the supports of the fourth-floor walkway were not designed to carry the second-floor walkway in addition to its own dead and live loads, so a collapse occurred, 114 people died, and millions of

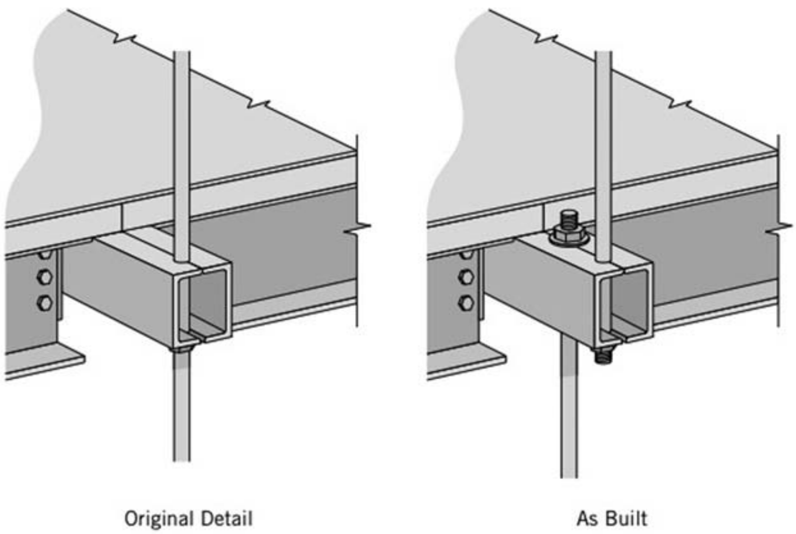



**Figure 1.4** The walkway suspension connection in Hyatt Regency Hotel in Kansas City, as originally designed and as built. The change made during construction left the second-floor walkway hanging from the fourth-floor walkway, rather than from the roof truss.

dollars of damage was sustained. If the fabricator had understood the designer's *intention* to hang the second-floor walkway directly from the roof truss, this accident might never have happened. Had there been a way for the designer to explicitly communicate his intentions to the fabricator, a great tragedy might have been avoided.

There's another lesson to be learned from the separation of the "making" from the "designing." If the designer had worked with a fabricator or a supplier of threaded rods while he was still designing, he would have learned that no one made threaded rod in the lengths needed to hang the second-floor walkway directly from the roof truss. Then the designer could have sought another solution in an early design stage. It was the case for many years that there was a "brick wall" between design engineers on one side and manufacturing engineers and fabricators on the other. Only recently has this wall been penetrated. Manufacturing and assembly considerations are increasingly addressed *during* the design process, rather than afterward. One element in this new practice is *design for manufacturing*, in which the ability to make or fabricate an artifact is specifically incorporated into the design requirements, perhaps as a set of manufacturing constraints. Clearly, the designer must be aware of parts that are difficult to make or of limitations on manufacturing processes as her design unfolds. The Hyatt Regency tale and the lessons drawn from it show us that communication is really important. Unless a design's fabrication specifications are complete and unambiguous, and unless they clearly convey a designer's intentions, the device or system won't be built in accord with the requirements set out by the designer. In short, design is a human activity, a social process. This means that *communication among and between stakeholders remains a preeminent, consistent, and ongoing concern*.

## 1.3 LEARNING AND DOING ENGINEERING DESIGN

Design is rewarding, exciting, fun, even exhilarating. But good design doesn't come easily. In fact, achieving excellence requires serious intellectual effort. That is why learning and doing (and teaching) design is challenging.

### 1.3.1 Engineering Design Problems are Challenging

Engineering design problems are challenging because they are usually *ill structured* and *open-ended*:

- Design problems are considered *ill structured* because their solutions cannot normally be found by applying mathematical formulas or algorithms in a routine or structured way. While mathematics is both useful and essential in engineering design, it is not possible to apply formulas to problems that are not well bounded or even defined. In the early stages of design, "formulas" are either unavailable or inapplicable. In fact, some experienced engineers find design difficult, simply because they can't fall back on structured, formulaic knowledge—but that's also what makes design a fascinating experience.
- Design problems are *open-ended* because they typically have several acceptable solutions. Uniqueness, so important in many mathematics and analysis problems,

© seraficus / iStockphoto © ZargonDesign / iStockphoto © Cora Reed / Shutterstock J. A. Kraulis / Masterfile

**Figure 1.5** A set of ladders that "enable people to reach heights they would be otherwise unable to reach" and suggest that design objectives involve more than just getting people up to some height.

simply does not apply to design solutions. In fact, more often than not, designers work to reduce or bound the number of design options they consider, lest they be overwhelmed by the possibilities.

Evidence for these two characterizations can be seen in the familiar ladder. Several ladders are shown in Figure 1.5, including a stepladder, an extension ladder, and a rope ladder. If we want to design a ladder, we can't even select a particular ladder type until we determine a specific set of uses for that ladder. Even if we decide that a particular form is appropriate, such as a stepladder, other questions arise: Should the ladder be made of wood, aluminum, plastic, or a composite material? How much should it cost? And, how much should the ladder support? Can we identify the *best* ladder design or the *optimal design*? The answer is, "No," we can't stipulate a ladder design that would be universally regarded as the best or that would be mathematically optimal in every dimension.

How do we talk about some of the design issues, for example, purpose, intended use, materials, cost, and possibly other concerns? In other words, how do we articulate the choices and the constraints for the ladder's form and function? There are different ways of representing these differing characteristics by using various "languages" or representations. But even the simple ladder design problem shows how the two characteristics of being open-ended (e.g., what kind of ladder?) and ill defined (e.g., is there a formula for ladders?) make design a difficult subject. How much more complicated and interesting are projects to design a new automobile, a skyscraper, or a way to land a person on Mars?

### 1.3.2 Learning Design by Doing

Teaching someone *how* to do design is not that simple. Like riding a bike, painting, or dancing, it often seems easier to tell a student, "*Watch* what I'm doing and then try to do it yourself." There is an element of *learning by doing*, which we call a *studio* aspect, in trying to teach any of these activities.

One of the reasons that it is hard to teach someone how to do design—or to throw a ball or draw or dance—is that people are often better at *demonstrating* a skill than they are at *articulating* what they know about applying their individual skills. Some of the skill sets just mentioned involve physical capabilities, but the difference of most interest to us is not simply that some people are more gifted physically than others. What is really interesting is that a talented softball pitcher cannot tell you just how much pressure she exerts when holding the ball, nor exactly how fast her hand ought to be going, or in what direction, when she releases it. Yet, somehow, almost by magic, the softball goes where it's supposed to go and winds up in the hands of a catcher. The real point is that the thrower's nervous system has somehow acquired the knowledge that allows her to assess distances and choose muscle contractions to produce a desired trajectory. While we can model that trajectory, given initial position and velocity, we do not have the ability to model the knowledge in the nervous system that generates that data. The pitcher has a combination of muscle memory, discipline, training, and practice that allows her to repeat the pitch time and again.

In a similar way designers, like dancers and athletes, *use drills and exercises* to perfect their skills, *rely on coaches* to help them improve both the mechanical and interpretive aspects of their work, and *pay close attention* to other skilled practitioners of their art. Indeed, one of the highest compliments paid to an athlete is to say that he or she is "a student of the game."

## 1.4 MANAGING ENGINEERING DESIGN PROJECTS

Good design doesn't just happen. Rather, it results from careful thought about what clients and users want, and about how to articulate and realize design requirements. That is why this book focuses on tools and techniques to assist the designer in this process. One particularly important element of doing good design is *managing* the design project. Just as thinking about design in a rigorous way does not imply a loss of creativity, using tools to manage the design process doesn't mean we sacrifice technical competency or inventiveness. On the contrary, there are many organizations that foster imaginative engineering design as an integral part of their management style. At 3M, for example, each of the more than 90 product divisions is expected to generate 30% of its annual revenues from products that didn't even exist five years earlier. So we will also introduce a few management tools that are useful in design projects.

We began this chapter by defining terms and developing a common vocabulary for design; we will do the same for management, project management, and the management of design projects in Chapter 16. For now, it will suffice to introduce the "3S model of project management." To be successful, a design project must track scope, schedule, and spending:

- **scope** *n*: deciding what a project must accomplish to be successful.
- **schedule** *n*: making sure that resources needed to accomplish the project scope are available and used when needed to complete the project by its agreed-upon due date.
- **spending** *n*: ensuring that a design project uses only the resources necessary to complete the project on time.

*Project management* is the tracking of these three matters to accomplish the goals and objectives of a project. All engineering design projects can be defined in terms of their goals, resources, and a need to finish in a fixed time frame. A number of tools have been developed to help project managers track a project's scope, scheduling, and spending. These include tools for understanding and listing the work to be done, scheduling the tasks to be done logically and efficiently, assigning tasks to individuals, and monitoring both project progress and expenditures. We will explore some of these tools as they are applicable to design projects in Chapter 16.

The precision in scope and spending spoken of in the context of project management may seem somewhat at odds with the open-ended nature of design. This is certainly the case when we try to predict the final form or outcome of a design project. Unlike a construction project, where the expected results are clearly articulated, a design project, especially a conceptual design project, may have a number of possible successful outcomes, or none! This makes the task and tools of project management only partially useful in design settings. As a result, we will present only project management tools that we have found to be useful in managing design projects conducted by small teams.

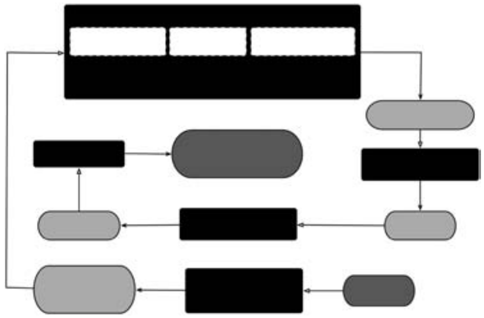
## 1.5 NOTES

*Section 1.2*: Our definition of engineering design draws heavily on Dym and Levitt (1991), Dym (1994), and Dym et al. (2005). Simon's definition of design is based on a set of lectures that were published as *The Sciences of the Artificial* (1981). The definition of engineering systems is taken from de Weck, Roos, and Magee (2011).

CHAPTER 2

# DEFINING A DESIGN PROCESS AND A CASE STUDY

*How do I do engineering design? Can you show me an example?*



HAVING DEFINED engineering design and some vocabulary, we now define a *process* of design, that is, how we actually do a design. This may seem a bit abstract, because we will break down a complex process into smaller, more detailed *design tasks*. However, as we define those design tasks, we will identify specific design tools and methods that we use to implement a design process. Keep in mind that we are ***not*** presenting a *recipe* for doing design. Instead, we are outlining a framework within which we can *articulate and* ***think*** about what we are doing as we design something. Further, it is important to keep in mind that our overall focus will be on what we will identify as *conceptual design*, the early stage where different design ideas or *concepts* are developed and analyzed.

## 2.1 THE DESIGN PROCESS AS A PROCESS OF QUESTIONING

Imagine you are working in a company that makes diverse consumer projects, and your boss calls you into her office and says, “Design a safe ladder.” You wonder to yourself: Why does anyone need still another ladder? Aren’t there a lot of safe ladders already on the market? And what does she mean by a “safe ladder”?

It's not a big surprise that a whole bunch of questions immediately come to mind. Typically, design projects start with a statement that talks about a client's intentions or goals, the design's form or shape, its purpose or function, and perhaps some things about legal requirements. That statement then leads to the designer's first task: to *clarify* what the client wants in order to translate those wishes into meaningful *objectives* (goals), *constraints* (limits), and *functions* (what the design has to do). This clarification task proceeds as the designer asks the client to be more precise about what she really wants.

Asking questions is an integral part of the design process. Aristotle noted long ago that *knowledge resides in the questions that can be asked and the answers that can be provided*. By looking at the kinds of questions that we can ask, we can articulate the design process as a series of *design tasks*. For example, with regard to designing a ladder, we

*establish a client's objectives* when we ask questions such as:

- Why do you want another ladder?
- How will the ladder be used?
- What market we are targeting?

*identify the constraints* that govern the design with questions such as:

- What does "safe" mean?
- What's the most you're willing to spend?

*establish functions* that the design must perform and suggest *means* by which those functions can be performed with questions such as:

- Can the ladder lean against a supporting surface?
- Must the ladder support someone carrying something?

*establish specifications* for the design with questions such as:

- How much weight should a safe ladder support?
- How high should someone on the ladder be able to reach?

*generate design alternatives* with questions such as:

- Could the ladder be a stepladder or an extension ladder?
- Could the ladder be made of wood, aluminum, or fiberglass?

*model* and *analyze* the design with questions such as:

- What is the maximum stress in a step supporting the "design load"?
- How does the bending deflection of a loaded step vary with the material of which the step is made?

*test* and *evaluate* the design with questions such as:

- Can someone on the ladder reach the specified height?
- Does the ladder meet OSHA's safety specification?

*refine* and *optimize* the design with questions such as:

- Are there other ways to connect the steps?
- Can the design be made with less material?

*document* the design process and *communicate* the completed design with questions such as:

- What is the justification for the design decisions that were made?
- What information does the client need to fabricate the design?

Thus, the questions we asked about the design establish steps in a process that move us from a problem statement through increasing levels of detail toward an engineering solution. The idea is to translate a client's wishes into a set of *specifications* that state in engineering terms how the design is to function or behave. These are benchmarks against which we can measure a design's performance.

With specifications in hand, we generate different *concepts* of how the design might work or look, that is, we create *design alternatives*. Then we choose one concept (say, a stepladder) and *build and analyze a model* of that concept, *test and evaluate* that design, *refine and optimize* some of its details, and then *document* the justification for the stepladder's final design and its fabrication specifications. In Section 2.2 we will present *all* of the tasks of the engineering design process in greater detail.

Some of the early clarification questions clearly connect to later tasks in the process. We make choices, analyze how competing choices interact, assess trade-offs in these choices, and evaluate the effect of these choices on our top-level goal of designing a safe ladder. For example, the ladder's *form* or shape and layout are strongly related to its *function*: We are more likely to use an extension ladder to rescue a cat from a tree and a stepladder to paint the walls of a room. Similarly, the weight of the ladder has an impact on how it can be used: Aluminum extension ladders have replaced wooden ones largely because they weigh less. The material of which a ladder is made affects not only its weight, but also its cost and its feel: Wooden extension ladders are both stiffer and heavier than their aluminum counterparts, so users of aluminum ladders feel a certain amount of "give" or flex in their lighter ladders.

Some of the questions in the later design tasks can be answered by applying mathematical models such as those used in physics. For example, Newton's equilibrium law and elementary statics can be used to analyze the stability of the ladder under given loads on a specified surface. We can use beam equations to calculate deflections and stresses in the steps as they bend under the given foot loads. *But there are no equations that define the meaning of "safe," or of the ladder's marketability, or that help us choose its color*. Since there are no equations for safety, marketability, color, or for many of the other issues in the ladder questions, we must find other ways to think about this design problem.

It is clear that we will face a vast array of choices as our design evolves. In our ladder design, we have to choose a *type* of ladder. We then have to decide how to fasten the steps to the ladder frame. These choices will be influenced by two things: (1) the desired behavior (e.g., although the ladder itself may flex, we don't want individual steps to have much give with respect to the ladder frame); and (2) manufacturing or assembly considerations (e.g., would it be better to nail in the steps of a wooden ladder, use dowels and glue, or nuts and bolts?). Note that we may decompose the ladder into its components to select among particular design choices.

As we work through these design questions and tasks, we are always communicating with others about the ladder and its various features. When we question our client about the ladder's desired properties, or the laboratory director about evaluation tests, or the manufacturing engineer about the feasibility of making certain parts, we are interpreting aspects of the ladder design in terms of *languages* and parameters that these experts use in their own work: We draw pictures in *graphical languages*; we write and apply formulas in the *language of mathematics*; we ask verbal questions and provide *verbal descriptions*; and we use *numbers* all of the time to fix limits, describe test results, and so on. Thus, the design process can't proceed without recognizing different design languages and their corresponding interpretations.

Our simple design problem illustrates how we might *formalize* the design process to make explicit the design tasks that we are doing. We are also *externalizing* aspects of the process, moving them from our heads into a variety of recognizable languages to be able to communicate with others. Thus, we learn two important lessons from our ladder design project:

- *The designer must fully understand what is needed from the final design.*
- *The designer must be able to translate the client's wishes into the languages of engineering design (e.g., words, pictures, numbers, rules, formulas, and properties) in order to model, analyze, test, evaluate, refine, optimize, and finally document the design.*

## 2.2 DESCRIBING AND PRESCRIBING A DESIGN PROCESS

We have just seen that asking increasingly detailed questions exposed several design tasks. We will now formalize such design tasks into a design process. Many design process models are *descriptive*: they *describe* the elements of the design process. Other models are *prescriptive*: they *prescribe* what must be done during the design process. We will first briefly present some descriptive models, and then introduce an extended set of our design tasks to convert one simple descriptive model into a more detailed prescriptive model.

The simplest descriptive model of the design process defines three phases:

1. *Generation*: the designer *generates* or creates various design concepts.
2. *Evaluation*: the designer *tests* the chosen design against metrics that reflect the client's objectives and against specifications that stipulate how the design must function.
3. *Communication*: the designer communicates the final design to the client and to manufacturers or fabricators.

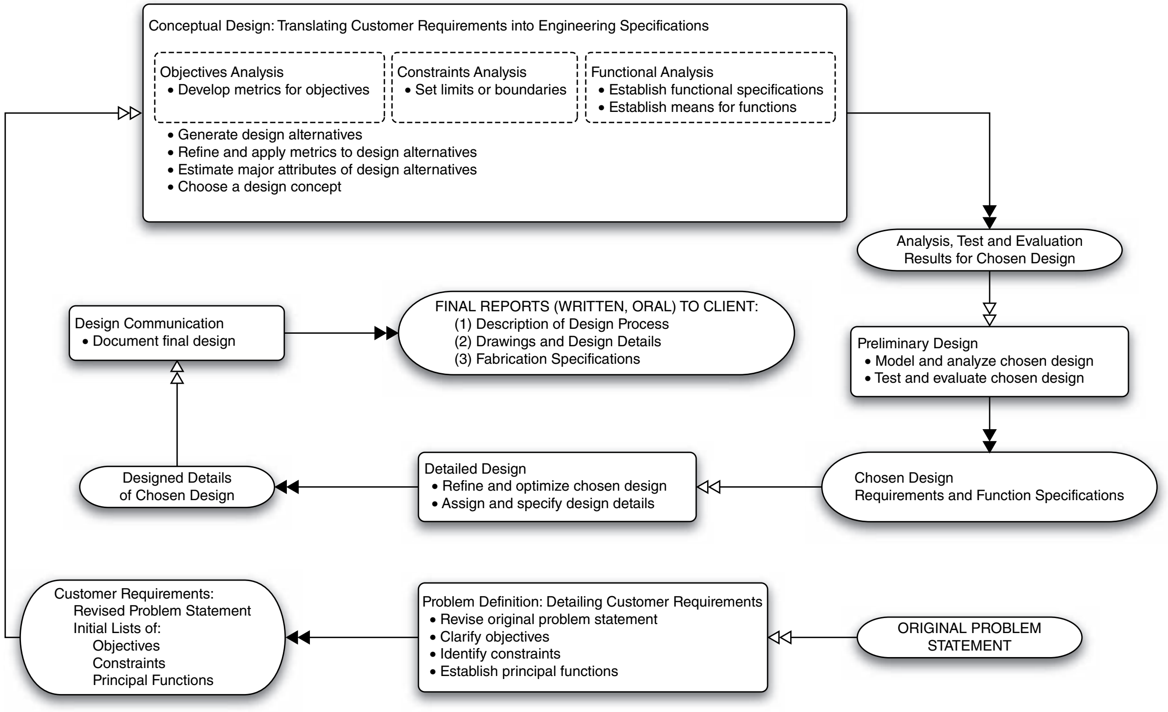


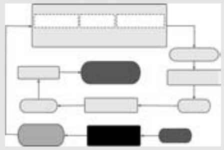

**Figure 2.1** A five-stage *prescriptive* model of the design process, presented as a *spiral* to convey the idea that design is not a simple linear sequence of tasks to be done. The design *stages* are in rectangles, and each stage's *outputs* are in ovals.

Another three-stage model splits up the design process differently: *Conceive*, *design*, and *implement* a final design, with the context providing meanings for these three steps. These two models are simple, but they are very abstract and provide no useful advice on *how* to actually generate designs.

We show a more extensive *prescriptive* model of the design process in Figure 2.1. It has five phases, shown in boxes with rounded corners, starting with an initial problem statement and ending with final design documentation. Figure 2.1 also shows, in ovals, the output of each design phase that also serves as input to the next design phase, and it displays the links between the five stages of this design process. We can also delineate the model and its design tasks in charts that describe each phase, showing the *input(s)* to that phase, the *design tasks* to be performed, and the *output(s)* or product(s) of that phase that in turn work as input to the next phase:

- *Problem definition*: We *frame the problem* by delineating the *customer requirements*, which means clarifying the client's objectives, identifying constraints, and establishing functions *before* we begin conceptual design.

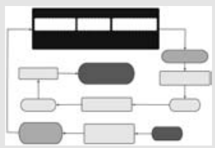
**1.** During *problem definition we frame the problem* by clarifying objectives, identifying constraints, establishing functions, and gathering the other information needed to develop an unambiguous statement of a client's wishes, needs, and limits, that is, the *customer requirements*.



Input: *original problem statement*
Tasks: *revise client's problem statement*
*clarify objectives*
*identify constraints*
*establish principal functions*
Outputs: *customer requirements:*
*revised problem statement*
*initial list of final objectives*
*initial list of constraints*
*initial list of principal functions*

- *Conceptual design*: We generate different *concepts* or *schemes* to achieve a client's objectives, satisfy constraints, and perform functions. Enough details (e.g., the spatial and structural relationships of the principal components) are worked out to estimate costs, weights, overall dimensions, and so on. Ladder concepts might be an extension ladder, a stepladder, or a rope ladder. We evaluate these concepts first translating the customer requirements (i.e., objectives, constraints, and functions) into *engineering specifications* that we use to articulate and benchmark our design.

**2.** In the *conceptual design* stage of the design process we translate the *customer requirements* into *engineering specifications* to generate *concepts* or *schemes* of *design alternatives* or feasible (i.e., acceptable) designs.



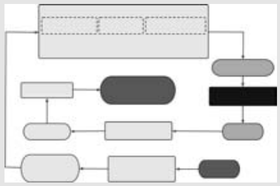
Input: *customer requirements*
*revised problem statement*
*initial list of final objectives*
*initial list of constraints*
*initial list of principal functions*
Tasks: *establish functional specifications*
*establish means for functions*
*write limits or boundaries of constraints*
*develop metrics for objectives*
*generate design alternatives*
*refine and apply metrics to design alternatives*
*estimate design alternatives' major attributes*
*choose a design concept*
Output: *a chosen design*
*analysis, test, and evaluation results for chosen design*

With its focus on trade-offs between high-level objectives, conceptual design is clearly the most abstract and open-ended part of the design process. Its output may include several competing concepts. Some argue that conceptual design *should* produce two or more schemes since early commitment to or fixation on a single design choice may be a mistake. This tendency is so well known among designers that it has produced a saying: "Don't marry your first design idea."

- *Preliminary design or embodiment of schemes*: Here we flesh out our proposed concepts, that is, we embody or endow design schemes with preliminary versions of their most important attributes. We select and size the major subsystems, based on lower-level concerns that take into account the performance and operating requirements. For a stepladder, for example, we size the side rails and the steps, and perhaps decide how to fasten the steps to the side rails.

**3.** In the *preliminary design* phase we identify and preliminarily size/estimate the principal attributes of the chosen design concept or scheme.

| | |
|---|---|
| Input: | *a chosen design* |
| | *specifications* |
| Tasks: | *model and analyze chosen design* |
| | *test and evaluate chosen design* |
| Output: | *analysis, testing, evaluation of chosen design* |

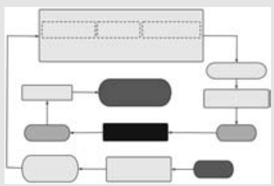


Preliminary design is definitely more "technical": We might do back-of-the-envelope or computer calculations. We make extensive use of rules of thumb about size, efficiency, and so on, that reflect our design experience.

- *Detailed design*: We now articulate our final design in much greater detail, refining the choices we made in preliminary design down to specific part types and dimensions. We use detailed design knowledge and procedures expressed in specific rules, formulas, and algorithms that are found in design codes (e.g., the ASME Pressure Vessel and Piping Code, the Universal Building Code), handbooks, databases, and catalogs.

**4.** During *detailed design* we refine and optimize the final design and assign and fix the design details.

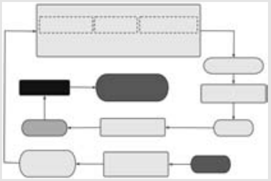
| | |
|---|---|
| Input: | *the analyzed, tested, evaluated design* |
| Tasks: | *refine, optimize the chosen design* |
| | *assign and specify the design details* |
| Output: | *proposed design and design details* |



- *Design communication*: We now spell out and present our design process, the resulting final design, and its fabrication specifications. In practice, the designer will usually have already developed much of the documentation along the way, and this communication phase will be more about tracking and organizing prior work products than writing a "new" report from "scratch."

**5.** Finally, during the *design communication* phase we document the fabrication specifications and their justification.

Input: *proposed design and design details*
Task: *document the final design*
Outputs: *final written, oral reports to client containing*:
*(1) description of design process*
*(2) drawings and design details*
*(3) fabrication specifications*



We have depicted the design process in Figure 2.1 as a *spiral* for several reasons. Design processes are often described as a linear sequence that seems to imply: do task 1, then do task 2, then do task 3. In practice, we actually keep completed phases and tasks in our minds as our design unfolds, and we refer back to them regularly. We may wonder while deep in a design project, "Why are we doing this?" By looking back at a project's objectives or constraints, or at a design decision we've already made, we can answer this question.

There are two other important elements that we hope the spiral depiction will help reinforce: *feedback* and *iteration*. Feedback occurs in two notable ways in the design process. First, *internal feedback* comes during the design process as test and evaluation results are used to *verify* that the design performs as intended. This feedback may come from the client and from internal customers, such as manufacturing (e.g., can it be made?) and maintenance (e.g., can it be fixed?). Second, *external feedback* comes after a design reaches its market and user feedback *validates* (or not) a successful design.

*Iteration* occurs when we repeatedly apply a common method or technique at different points in a design process. For example, we might write equilibrium equations for an entire structure, and then at a lower *level of abstraction* (i.e., on a different scale) we write equilibrium equations for structural components). Similarly, as we fix more details in a design and become less abstract, and we might review and reestablish means for our functions. Again, we always want to keep in mind the original objectives, constraints, and functions as we get closer to our final design. This may also mean that we have to do some redesign, in which case we will certainly repeat tasks such as analyzing the design or testing and evaluating the design.

Given that there are feedback loops and that we will reiterate some tasks, why didn't we include them in Figure 2.1? As important as feedback and iteration are, it is also important not to be overly distracted by these adaptive characteristics when learning about and doing design for the first time.

We now have a “checklist” we can use to ensure that we have done all of the “required” steps. Lists like this are often used by design organizations to specify and propagate approaches to design within their firms. However, we should keep in mind that this and other detailed elaborations add to our understanding of the design process only in a limited way. At the heart of the matter is our ability to model the tasks done within each phase of the design process. With this in mind, we will present some means and formal methods for doing these design tasks in Section 2.3.

## 2.3 INFORMING A DESIGN PROCESS

Even prescriptive descriptions of design processes can fail us because they don’t tell us *how* to generate or create designs, or even how to do some of the tasks specified. In this section we describe some ways of thinking about design, after which we list some of the formal design methods that we explore in greater detail in later chapters. We then list some of the kinds and sources of design knowledge that informs what we do as designers, including: how we acquire information; how we analyze information and test outcomes against desired results; and how we get feedback from clients, users, and other interested parties.

### 2.3.1 Informing a Design Process by Thinking Strategically

Strategies are ways of thinking about a problem or situation. Effective designers have habits of thought they bring to their work that help them realize better decisions. *Least commitment* is one general strategy for thinking about design: Don’t make decisions before you have to. This is a good habit of thought that guards against making decisions before there is a reason to make them. Premature commitments can be dangerous because we might become attached to a bad concept or we might limit ourselves to a suboptimal range of design choices. Least commitment is of particular importance in conceptual design because the consequences of any early design decision are likely to be propagated far down the line. It is generally unwise to commit to a particular concept or configuration until we are forced to because we’ve exhausted our information or range of choices or time available.

*Decomposition*, also known as *divide and conquer*, is another important habit of good design thinking: Break down, subdivide, or decompose larger problems into smaller subproblems. These smaller subproblems are usually easier to solve or otherwise handle. We do have to keep in mind that subproblems can interact, so we must ensure that the solutions to particular subproblems do not violate the assumptions or constraints of complementary subproblems.

### 2.3.2 Informing a Design Process with Formal Design Methods

Usually, when we think systematically about doing things, we can develop tools and techniques we can use. We now present a brief introduction to the formal design methods identified in the charts by introducing some of the formal tools we will use in following the design process.

*Objectives trees* are hierarchical lists of the client's objectives or goals for the design that branch out into tree-like structures. We build objectives trees in order to clarify and better understand a client's project statement. The objectives that designs must attain are clustered by sub-objectives and then ordered by degrees of further detail. For example, we might take an objective like "portable," and break that down into subobjectives: "lightweight" and "small when collapsed for storage." The highest level of abstraction of an objectives tree is the top-level design goal, derived from the client's project statement. In Chapter 4 we explain how to construct objectives trees and explore the kinds of information we learn from them.

*Pairwise comparison charts* (PCCs) are used to rank order our design objectives. It is helpful, for example, to know whether it is more important to the client that our ladder is "portable" or "inexpensive" in case trade-offs need to be made. A PCC is a relatively simple device in which we list the objectives as both rows and columns in a matrix or chart and then compare them on a pair-by-pair basis, proceeding in a row-by-row fashion. We use PCCs early in the design process, and they are described in Section 4.4.

*Metrics* are created to measure how well we achieve a design's objectives, thus allowing us to evaluate design alternatives in terms of attributes the client desires. We explain how we create metrics in Section 4.5.

*Functional analysis* is used to identify what a design must do. Identifying and performing functions are central to engineering design, and we will develop several approaches (e.g., black and transparent boxes, dissection, enumeration, and function–means trees) to functional analysis in Chapter 6.

The *performance specification method* provides support for the elaboration of the *specifications* that reflect, in engineering terms, how a design will function. The aim is to list solution-independent attributes and performance specifications (i.e., "hard numbers") that specify the requirements of a design concept. We describe performance specifications (requirements) and their role in Section 6.2.

*Morphological charts* are used to identify the ways or means that can be used to make function(s) happen. Such "morph charts" express *functions* as *verb–noun action pairs*; the *means* are specific ways to use or convert energy, or to process information and/or materials. For example, if one of our ladder functions is to stabilize the user over uneven terrain, we could realize that by having adjustable feet, or by having a wide base. The morph chart provides a framework of the *design space*, an imaginary "space" that we can use to generate potential design alternatives for a design problem. We describe morph charts in Chapter 6.

### 2.3.3 Acquiring Design Knowledge to Inform a Design Process

The *literature review* is the classic way to find examples of prior work and determine the state of the art. We need literature reviews early on, in the framing and conceptual stages, to better understand the client, potential users, and the design problem itself. We should consider existing solutions, because there's little point or profit in reinventing the proverbial wheel. In detailed design, we may want to review available off-the-shelf parts and materials to help standardize our design and reduce fabrication costs. So at various stages we should look at the technical literature, patent listings, vendor literature, handbooks, material properties tables, and design and legal codes.

*Benchmarking* competitive products means evaluating the functionality and behavior of similar products already on the market. Benchmarking is often done to set a bar for a *better* product.

*Reverse engineering* or *dissection* is also about "seeing what's out there": it consists of dissecting or taking apart competitive or similar products. It is often done to assess functional behavior with the aim of developing better means to accomplish the same or similar functions.

*Informal interviews* of potential users should be undertaken very early in a design project to assist in problem definition. While informal interviews are relatively easy to conduct, it is important to be sensitive to the time and other constraints of the interviewees by preparing for the interviews, for example, provide the interviewees with advance lists of the topics and questions, and complete a significant part of the literature research *before* conducting those interviews.

*User surveys and questionnaires* are used in market research to identify users' views of the problem space and their response to possible solutions. Market research can help clarify a design problem in its early stages, especially with open-ended questions. Later surveys can be used together with PCCs and morph charts to help select a final design.

A *structured interview* melds the consistency of surveys with the flexibility of informal interviews by using a previously defined set of questions that may or may not be made available to the interviewees. An interviewer can follow up a particular response and open up new areas. A structured set of questions also assures the interviewee that the interview has both purpose and focus, and it provides an agenda that ensures that key matters will be covered.

*Focus groups* are an expensive way to elicit the response of appropriately selected users and others to potential designs. They are not often used by student design teams because they demand considerable sophistication in psychological matters and are expensive.

Structured *brainstorming* by a design team can also generate ideas and insights, and in opening up new avenues for research and analysis. However, productive brainstorming is a complex activity that requires thought, preparation, and appropriate professional behavior.

### 2.3.4 Informing a Design Process with Analysis and Testing

We can't know whether or not a design concept might work unless we can measure outcomes, and we can't measure outcomes without having a *metric* or a standard against which to measure those outcomes. While metric is a general term meaning a frame or ruler for measurement, we will specifically use the term metrics to measure how well a design's objectives are achieved. We will describe how we develop metrics in Section 4.4. We will also develop another set of measurements, *design specifications*, to state in engineering terms the functional performance of a design. This is an extremely important topic that we will discuss in depth in Sections 6.2 and 7.3.

*Experiments* and *testing* are often used to get data, in the field or in a lab, about how well potential design actually works. Testing can run the gamut from component testing to *proof-of-concept testing*, to *prototype development*, *and testing*. We will say a good bit more about such testing in Chapter 11.

In many cases we don't develop or test a prototype, perhaps due to cost, size, or hazards. In such cases, we may resort to *simulation* in which we exercise an analytical or computer model of a proposed design to simulate its performance under a stated set of conditions. However, we can only do that if we really understand the device we're modeling and its true operating conditions. With such deep understanding, we can get useful data from simulation that will enable us to evaluate against constraints, specifications, and any applicable standards. One outstanding example of such simulation is the use of wind tunnels and related computer analyses to assess the effects of wind loading on tall buildings, on long, slender suspension bridges and, of course, airplanes.

*Computer analysis* is closely related to simulation and uses computer-based models, typically discipline-specific, to analyze design components in preliminary and detailed design. Such computer analyses include finite element analysis, integrated circuit modeling, failure mode analysis, and criticality analysis.

### 2.3.5 Getting Feedback to Inform a Design Process

We noted earlier that feedback was an intimate part of design: We use internal feedback to verify that we're solving our design problem correctly, and external feedback to validate that our design has solved the right problem.

*Regularly scheduled meetings*, at which the progress of the design project is tracked and discussed, may be the most important means for obtaining feedback from clients and other members of the design team.

*Formal design reviews*, held at specified intervals, are a standard "best practice." We use them to update the client (and sometimes others) on the design status, typically including sufficient technical detail that the implications of the design can be explored and assessed. Design reviews involve a lot of "give and take" between the design team and its audience, and thus may seem harsh to young designers. A design team usually benefits by having to justify various technical details to clients and outside experts because its members become more aware of implicit unwarranted assumptions and errors or oversights.

In some design environments, *public hearings* are required by relevant civil laws or public policies in order to subject a design to public review and comment. Public hearings and meetings are increasingly the norm for major design projects (e.g., transportation or power projects), even when the client is a private entity.

We have already noted that *focus groups* are important sources of user input for problem definition. Such groups are also widely used to assess user reaction to designs as they near adoption and marketing.

Similarly, in industries like software design, an almost-but-not-quite-finished version of a product is released to a small number of users for *beta testing*. Beta tests allow designers to expose design or implementation errors and to get feedback about their product before it reaches a larger market.

## 2.4 CASE STUDY: DESIGN OF A STABILIZER FOR MICROLARYNGEAL SURGERY

One of the founding fathers of Harvey Mudd's design-intensive engineering program, the late Jack Alford, said that learning to design was like to learning to dance: "You have to get out on the dance floor and get your toes stepped on." In other words, *design is best learned* *by doing*. In that context, some design tools can be learned by doing exercises, and some by seeing how others use them. We now present a design case study and begin two design examples to illustrate how engineers do conceptual design.

The case study and one of the design examples are drawn from student work done by engineering majors in their first engineering course at Harvey Mudd College, *E4: Introduction to Engineering Design*. Student teams are assigned the task of developing and prototyping a conceptual design of a device or system. The projects are virtually always done for nonprofits or educational institutions, and they provide HMC students with the insight that good engineering design may be done in nontraditional, noncorporate settings. The course stresses the formal design methods we present in this book.

The design project we describe was undertaken by four teams of students on behalf of Brian Wong, M.D., of the Beckman Laser Institute of the University of California Irvine. The particulars of this case study are an edited combination of results obtained by the four teams that shows how design teams thought through the design process while they were designing a device for a client. The project was concerned with laryngeal surgery.

Laryngeal, or vocal cord, surgery is often required to remove growths such as polyps or cancerous tumors. The "lead" cells of such growths must be removed accurately and completely. Patients also incur the risk of damage to their vocal cords—and so their speech—during these surgeries. In spite of many other surgical advances over recent decades, laryngeal surgery has not changed much. One change that has occurred, however, is that surgeons now access the vocal cords through the mouth, rather than by cutting open the throat. This has made it harder to insert and stabilize both optical devices and surgical instruments that cut, suck, grasp, move and suture. Surgeons must be able to control their own tremors in order make accurate and precise cuts during the procedure.

Tremor is the natural, small-scale shaking of the hand. (Watch the movement of your own fingertips as you hold your hands straight out in front of you.) In the context of laryngeal surgery, such tremors tend to be amplified as the surgeons insert and control foot-long instruments in the patient's throat.

The project began when Dr. Wong presented the following *initial problem statement* to the student teams:

> *Surgeons who perform vocal cord surgery currently use microlaryngeal instruments, which must be used at a distance of some 12–14 in. to operate on surfaces with very small structure (1–2 mm). The tremor in the surgeon's hand can become quite problematic at this scale. A mechanical system to stabilize the surgical instruments is required. The stabilization system must not compromise the visualization of the vocal cords.*

The teams began by *defining* or *framing* this design problem. They talked with Dr. Wong and other physicians and did some basic library research to develop more information about laryngeal surgery. They learned that the abnormalities that were operated on were typically 1–2 mm wide, while the vocal cords themselves are approximately 0.15 mm thick. This meant that the physiological surgical tremors of the surgeon's hands had to be reduced from 0.5–3.0 mm to an acceptable tremor amplitude of 0.1 mm. They also learned that the surgeons needed to control the instruments at distances far from the patient's mouth (and vocal cords). One of the teams wrote the following *revised problem statement*:

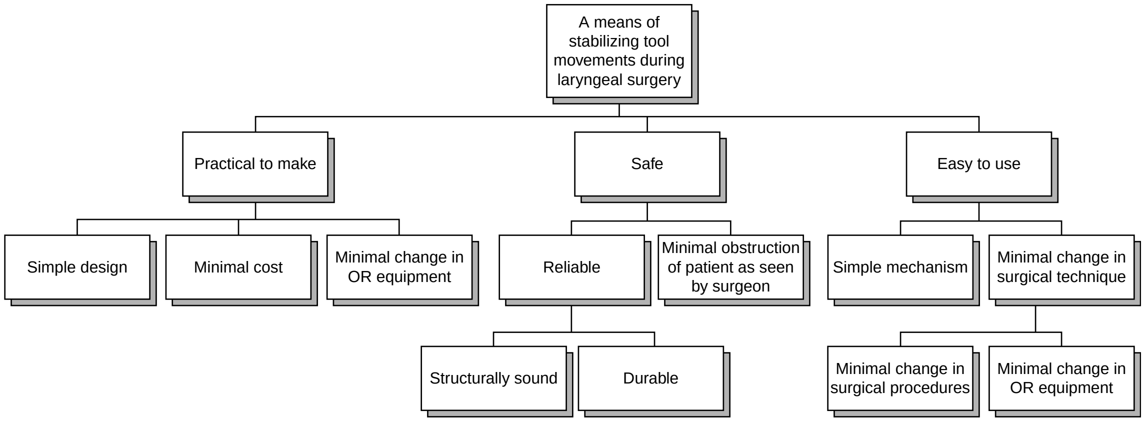



**Figure 2.2** The *objectives tree* displaying the client's objectives and subobjectives for the microlaryngeal stabilization device whose design is presented as a case study in Section 2.5.1. This tree is developed largely from the work of one of the three teams who worked on this project. After Chan et al. (2000).

*Microlaryngeal surgery seeks to correct abnormalities in the vocal cords. The abnormalities, such as tumors and cysts, are often 1–2 mm in size and are typically removed from the vocal cords, which are only 0.15 mm in size. During the operation, the surgeon must control his or her surgical instruments from a distance of 300–360 mm (12–14 in.) due to the difficulties in accessing the vocal cords. At this small scale, the physiological tremor in the surgeon's hand can be problematic. Design a solution that minimizes the effects of hand tremors in order to reduce unintentional movements at the distal end of the instrument to an amplitude of no more than 1/10 of a millimeter. The solution must not compromise visualization of the vocal cords.*

Note that this revised problem statement contains more detail and also excludes an implied "mechanical" solution referenced in the original problem statement.

Continuing with problem framing, the teams listed the client's *objectives* for the designed stabilizing device. Objectives and subobjectives are routinely displayed in an *objectives tree*. One team's objectives tree for this project is displayed in Figure 2.2. Two of the objectives are that the device should minimize obstruction of the surgeon's vision, and the manufacturing cost should be minimized. The teams also developed lists of *constraints*, that is, the strict limits that govern acceptable designs. A constraint list for the device includes the following:

- it must be made of nontoxic materials;
- it must be made of materials that do not corrode;
- it must withstand sterilization;
- its cost must not exceed $5000;
- it must not have sharp edges;
- it must not pinch or gouge the patient; and
- it must be unbreakable during normal surgical procedures.

**TABLE 2.1 A *pairwise comparison chart* created by one of the student teams to compare objectives for the microlaryngeal stabilization device. An entry "1" indicates the objective in that row is more important than that of the column in which it is entered. It shows that the reduction of the surgeon's tremor is the most important objective for this project**

| Goals | Reduce Tremor | Sturdy | Safe | Inexpensive | Easily Used | Score |
|---|---|---|---|---|---|---|
| Reduce Tremor | •••• | 1 | 1 | 1 | 1 | 4 |
| Sturdy | 0 | •••• | 0 | 1 | 0 | 1 |
| Safe | 0 | 1 | •••• | 1 | 1 | 3 |
| Inexpensive | 0 | 0 | 0 | •••• | 0 | 0 |
| Easily Used | 0 | 1 | 0 | 1 | •••• | 2 |

After Both et al. (2000).

We see that there is an upper limit on the cost and on the device having sharp edges, among others.

Another facet of framing the problem involved rank ordering the design objectives in terms of their perceived relative importance. This ranking was done using a PCC, which is an extension of what people do when comparing two objects, one against the other. The PCC, which will be explained in detail in Section 4.3, enables each objective to be compared to every one of the other objectives, The PCC produced by one of the teams is shown in Table 2.1; it shows that reducing the surgeon's tremor is the most important objective, while cost is the least important. This ranking helped focus the team's attention, as well as seeming to accord with our intuitions.

The teams then set about establishing *metrics* that would enable them (later in the design process) to measure whether various designs would achieve the objectives set out for the project. The metrics for two of the objectives of Figure 2.2, along with their units and scales, are

Objective: Minimize viewing obstruction.

Units: Rating percentage of view blocked on a scale from 1 (worst) to 10 (best).

Metric: Measure the percentage of view blocked by the instrument. On a linear scale from 1 (100%) to 10 (0%), assign ratings to the percentage of view blocked.

Objective: Minimize the cost.

Units: Rating cost on a scale of 1 (worst) to 5 (best).

Metric: Determine a bill of materials. Estimate labor, overhead, and indirect costs. Calculate the total cost. On a scale from 1 (worst) to 5 (best), assign ratings to the calculated cost as follows:

| Cost ($) | Points |
|---|---|
| 4000–5000 | 1 |
| 3000–4000 | 2 |
| 2000–3000 | 3 |
| 1000–2000 | 4 |
| 1–1000 | 5 |

Continuing on with their conceptual design tasks, the design teams turned to *functional analysis*, that is, to determining what a successful design will actually do. The teams determined the specific *functions* that their proposed devices must perform, along with *specifications*, which are the engineering expressions of the performance of the functions. The teams identified the required functions by applying some of the tools that will be discussed in detail in Section 6.1, including the *black box*, the *glass box*, and the *functions–means tree*. One team's list of functions states that the microlaryngeal stabilizer must do the following:

- stabilize the instrument;
- move the instrument;
- stabilize the distal end of the instrument;
- reduce surgeon's muscle tension (shaking tremors) during surgery; and
- stabilize itself.

The specification for the first of these functions was written as follows:

*Function*: Stabilize the instrument.

*Specification*: This function is not achieved if the design cannot reduce the amplitude of a trembling hand to less than 0.5 mm; it is optimally achieved if it controls the amplitude of a trembling hard to make it less than 0.05 mm; and it is overly restrictive if it inhibits or disallows any instrument or hand use.

With the functions and their specifications now largely established, the design process turned to *creating* or *generating alternative designs*. One excellent way to begin creating designs is to list each of the required functions in the left-hand column of a matrix, and then list across each functional row the various *means* by which each function can be implemented. The resulting matrix or chart is called a *morphological chart* or "morph chart." Figure 2.3 is a morph chart for the microlaryngeal stabilizer. Such morph charts effectively tell us how large a *design space* we are working in because each design candidate must achieve every function, no matter which of its implementations or means are used. Thus, for each of the six functions displayed in Figure 2.3, we select one of its following means to produce a possible design. As we will explain in more detail in Chapter 6, a given function's means may not necessarily connect with all of the means of all of the other functions, in which case there will be combinations that must be excluded. Still, the combinatorial effects can be daunting for designs that have many functions since each function can be implemented by several different means.

For the laryngeal stabilizer project, one design alternative positioned the surgical instrument at the end of a *lever*. A second design alternative had the instrument supported on a *stand*, moved by a system of *pulleys*, and supported by the *stand* itself. The instrument stand removes the need for surgeon to operate in a fixed position, which thus reduces tremor-inducing muscle tension. In the third design alternative, the surgeon's *hands* hold and move the instrument. Distal support is provided by *crosswires* attached directly to the laryngoscope. A *forearm rest* reduces the surgeon's tremor-inducing muscle tension. Figure 2.4 shows concept drawings for these three design alternatives.

| FUNCTIONS | POSSIBLE MEANS | | | | | | |
|---|---|---|---|---|---|---|---|
| Stabilize instrument | Hand | Stand | Clamp | Magnet | Edge of Laryngoscope | Wire | |
| Move instrument | Hand | Gears | Pneumatics | Ball bearings | Lever | Pulley | |
| Stabilize distal end of instrument | Magnet | Crosswires | Track System | Spring | Gyroscopes | Ball Bearings | Stand |
| Reduce muscle tension of surgeon during surgery | Instrument Stand | Hand Platform | Pillow | Elbow Platform | Forearm Rest | Shoulder Sling | |
| Stabilize itself | Gyroscope | Springs System | Stand | Magnet | Suspension System | Rest against stable surface | Attach to Laryngoscope |

**Figure 2.3** A *morphological chart* for the microlaryngeal stabilization device showing the *functions* and the corresponding *means* or *implementations* for each function. Possible design solutions are assembled by selecting one means from each row. After Chan et al. (2000).

We complete the conceptual design phase of the design process by narrowing the field of possible designs and, eventually, *selecting a final design*. We can do this by exercising a *decision* or *selection matrix* in which we evaluate each design based on how well it achieves each of the design objectives as measured by the metrics described just above. As we will point out in Chapters 4 and 8, such decision matrices must be used with care because the numbers are often subjective, for example, some of the metrics may be qualitative assessments rather than quantitative measurements. We cannot weight or add the scores for the objectives: we can only rank them in perceived order of importance. In the present case, for example, the design that was finally chosen by the client, the *crosswire* support would not have the highest total value, if we had summed the scores (see Figure 2.5).

Design concepts or ideas must also be tested in meaningful ways in order to ensure that they work. One of the student design teams tested its concept by attaching a pencil lead to the end of a surgical tool and traced a predrawn square with the instrument both with and without their design being attached. As we can see from their test results, displayed in Figure 2.6, their designed device removed almost all of the tremor.

Finally, after all of the work and the selections, the design process is complete and a design can be documented and turned over to the client. In this case, one of the selected designs is being used in laryngeal surgeries and being prepared for manufacture as a medical product.

This case study has presented many of the design tools that are the main focus of this book. We have not shown the management tools used by the teams, although the design teams on this project did use them, and we have said nothing about the dynamics of each of the three design teams. These as yet unsung elements are also very important for achieving effective design results.

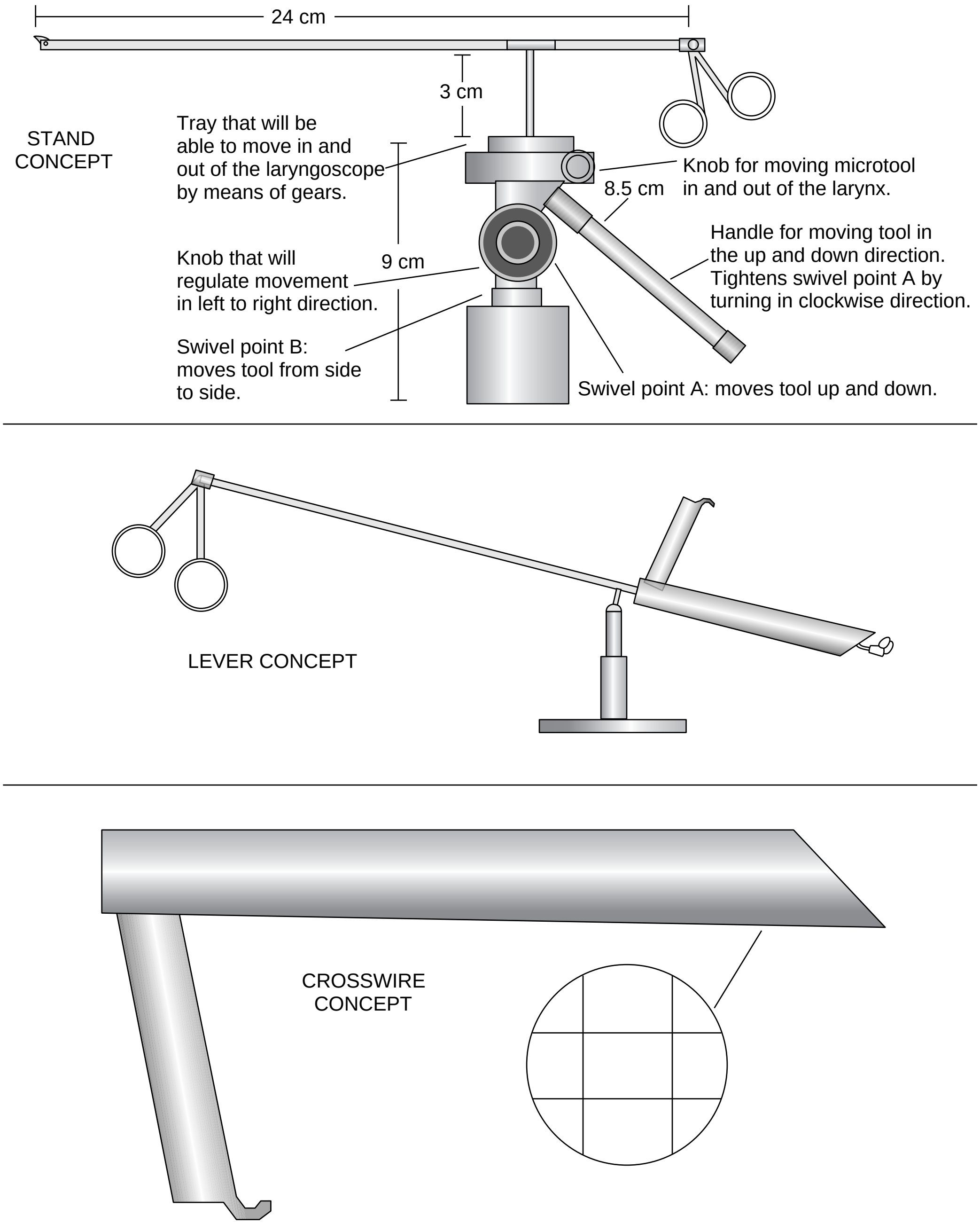



**Figure 2.4** Three *design alternatives* produced by the student teams that worked on the microlaryngeal stabilization device for surgeons at the University of California at Irvine. Dr. Wong and his colleagues have adopted the crosswire concept for clinical trials. After Both et al. (2000), Chan et al. (2000), and Saravanos et al. (2000).

| DESIGN CONSTRAINTS | LEVER | INSTRUMENT STAND | LARYNGOSCOPE CROSSWIRES |
|---|---|---|---|
| C: Must not fall apart during surgery | y | y | y |
| C: Noncorrosive materials | y | y | y |
| C: Must withstand medical sterilization procedures (autoclave, enzymatic, bleach, etc.) | y | y | y |
| C: Cannot get in the way of surgical instruments | y | y | y |
| C: Cannot block the view of vocal cords | y | y | y |
| C: Materials must be compatible with human body | y | y | y |
| C: Must be easy to clean with conventional means (scrub brush, water jet, soaking, etc.) | y | y | y |
| C: Cannot cost more than $5000 | y | y | y |
| DESIGN OBJECTIVES | Score | Score | Score |
| O: Structurally sound | 75 | 85 | 80 |
| O: Strong materials | 85 | 90 | 85 |
| O: Minimum obstruction of vocal cords | 100 | 100 | 65 |
| O: Minimum obstruction between patient and surgeon/nurse | 65 | 70 | 100 |
| O: Simple design | 60 | 70 | 90 |
| O: Minimum cost | 50 | 70 | 55 |
| O: Compatible with existing instruments | 30 | 80 | 50 |
| O: Minimal alteration of existing surgical procedures | 45 | 80 | 70 |
| O: Compatible with existing instruments | 30 | 85 | 80 |
| O: Simple mechanism | 70 | 60 | 95 |

**Figure 2.5** A *decision* or *selection matrix* used by one of the student teams that worked on the microlaryngeal stabilization device to select a final design. The decision matrix, whose numbers should be taken with caution, suggests which designs are preferred. After Chan et al. (2000).

## 2.5 ILLUSTRATIVE DESIGN EXAMPLES

We now briefly describe two illustrative examples that will be carried through the remainder of the book to illustrate the various design techniques. The first illustrative example is the design of a container for a new juice product. A summary of the juice container project is:

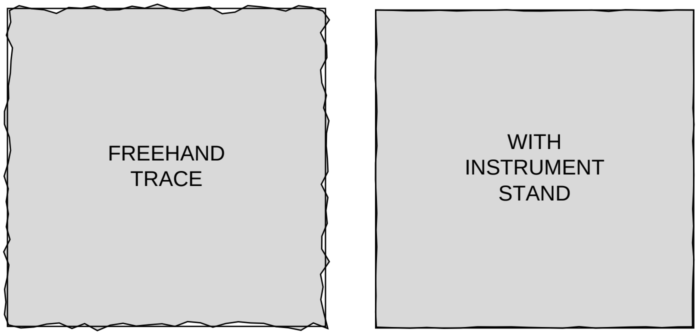



**Figure 2.6** An example of the *testing* done by one student team to show that their concept successfully stabilized the surgeon's hand and reduced tremor, as demonstrated by the successful tracing of a predrawn square. After Chan et al. (2000).

Design of a container to deliver a new children's juice. *This is a stylized industrial design project that highlights some of the early questions that must be addressed before a designer can apply more conventional engineering science knowledge to the problem.*

*Designers: Dym, Little, and Orwin LLC*
*Clients: American Beverage Company (ABC) and National Beverage Company (NBC)*
*Users: Children living both in the United States and abroad*
*Initial Problem Statement: Design a bottle for a new children's juice product.*

The second illustrative design example is based on the work done by students in the first-year design course at Harvey Mudd College. We use their results to illustrate and further explain how formal design methods are used. In many cases, the design teams' results are displayed exactly as they presented them in their final reports at the end of the semester, without any further edits. This second illustrative example project is:

Design of an arm restraint device for children with cerebral palsy. *An arm restraint was designed by a team of students in Harvey Mudd College's first-year design course.*

*Designers: Teams of students in HMC's first-year design class*
*Client: Danbury School, Claremont, California*
*Users: Students at Danbury School diagnosed with cerebral palsy (CP)*
*Abbreviated Problem Statement: Design a device to stabilize the arm of a student and counteract her involuntary, CP-induced tremors as she writes or draws.*

## 2.6 NOTES

*Section 2.1*: The stepladder example derives from the first-year design course taught at Harvey Mudd College and is briefly described in Dym (1994b). The paraphrasing of Aristotle's observation is from Dym et al. (2005).

*Section 2.2*: As with definitions of design, there many descriptions of the design process, and many of them can be found in Cross (1994), Dym (1994a), French (1985, 1992), Pahl and Beitz (1996), and VDI (1987). Further descriptions of the tasks of design can be found in Asimow (1962), Dym and Levitt

(1991), and Jones (1992). Examples of the application of conceptual design tools as problem-solving tools can be found in Schroeder (1998) for automobile evaluation and in Kaminski (1996) for college selection. The create–design–implement abstraction is also the basis for the model CDIO engineering curriculum (Crawley et al. 2007).

*Section 2.3*: Further elaboration of strategic thinking in design appears in Dym and Levitt (1991). More elaborate descriptions of formal design methods can be found in Cross (1989), Dym (1994a), French (1985, 1992), Pahl and Beitz (1984), and VDI (1987). Detailed descriptions of ways to acquire and process knowledge are given in Bovee, Houston, and Thill (1995), Ulrich and Eppinger (2000), and Jones (1992).

*Section 2.4*: The microlaryngeal stabilization device case study is detailed in Both et al. (2000), Chan et al. (2000), Feagan et al. (2000), and Saravanos et al. (2000).

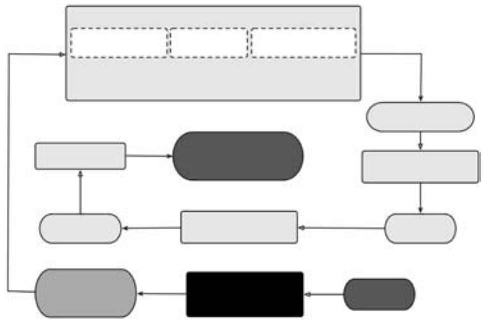
PART II

# THE DESIGN PROCESS AND DESIGN TOOLS

CHAPTER 3

# PROBLEM DEFINITION: DETAILING CUSTOMER REQUIREMENTS

*What does the client require of this design?*



IN CHAPTERS 1 and 2, we defined engineering design and described a process for doing design. Now we turn to *problem definition*, when we try to understand what the client *requires* in this design by casting the design problem in terms of *objectives* to be achieved, *constraints* that must be satisfied, and *functions* the design must perform. As our starting point we closely examine the design problem as stated by the client, in order to make us aware of the design space in which we will be working.

Recall that in Figure 2.1, revising the problem statement, clarifying objectives, identifying constraints, and establishing functions all appeared in the same "block," which is the highlighted black element on the design icon just above. This is no accident. Designers generally review prior art, confer with clients and users, and speculate among themselves about all of these important intellectual objects simultaneously. This allows the designers to efficiently gain information from their sources, but this information must ultimately be organized into a framework. In this chapter, we look at how we start on the tasks of problem definition, beginning with the initial statement from the client. We then present ways to organize all of the information found from our design inquiries, and finally develop a revised problem statement, which is one of the outputs of our problem definition box. To be successful in the problem definition stage, we will need to examine objectives, constraints, and functions in detail; we do that in the three chapters that follow.

## 3.1 CLARIFYING THE INITIAL PROBLEM STATEMENT

Most design projects begin when a *client* sets out a *problem* to be solved, typically in a verbal *problem statement* that identifies a gadget that will appeal to certain markets (e.g., a container for a new drink), a widget that will perform some specific functions (e.g., a chicken coop), or a problem to be fixed through a new design (e.g., a new transportation network and hub). The initial problem statement may be short. Imagine working for a food company whose management challenges its designers: "Design a bottle for our new children's fruit juice product." One response to this challenge is to design a clever new label for an existing bottle and declare the work done. But, is this a *good* design? Is it the *right* design? There is no way to answer these questions because the problem statement is so brief that it gives no hint of other considerations that might enter into thinking about or assessing the design, for example, the intended market, the shape or materials choice of the container, and so on. Indeed, we might need more information on the nature of the juice product to be confident we had a safe design.

Another problem statement might take a longer form: "The Claremont Colleges need to reconfigure the intersection of Foothill Avenue and Dartmouth Avenue so students can cross the road." This communicates someone's idea of what the problem is, but it may also contain errors, show biases, or imply solutions. *Errors* may include incorrect information, faulty or incomplete data, or simple mistakes regarding the nature of the problem. In this instance, the error is minor—Foothill *Boulevard*, not Foothill Avenue, bypasses the Claremont Colleges—but this could be confusing to someone unfamiliar with Claremont. *Biases* are presumptions about the situation that may also prove inaccurate because the client or the users may not fully grasp the entire situation. In our traffic example, for instance, the problem may not be related to the design of the intersection at all but to the tendency of students to jaywalk or to a bias toward thinking in "concrete" terms that may affect our design process. *Implied solutions* (i.e., a client's best guess at the answer) frequently appear in problem statements. In the fruit juice project, for example, the management called for a bottle, while many other solutions, such as a carton or bag, are possible. In other cases, the implied solution may be more subtle, such as calling for a device when a process might be more appropriate. While implied solutions offer some insight into what a client is thinking, they restrict the design space in which a designer searches for a solution. Also, an implied solution may not solve the problem at hand. For example, the traffic problem may not be solved by reconfiguring the intersection. If students jaywalk, expensive changes to the intersection may do little or nothing to mitigate this. If the problem is that students are crossing a dangerous street, perhaps the destination to which they are headed should be relocated. The point is that we must carefully examine initial problem statements in order to identify and deal with errors, biases, and implied solutions. Only then can we begin to understand and solve the real problem.

We may want to do a quick rewrite of the initial problem statement that eliminates errors, biases, and implied solutions. We can then use that as a basis for discussing the project with our client, and as a guide along our design path. Some teams do such a revision of the problem statement very early in the design process, and then do another, more extensive one when they have come to an understanding of the objectives, constraints, and functions that define the design space. The latter version is discussed in Section 3.3.

## 3.2 FRAMING CUSTOMER REQUIREMENTS

We focus on getting a clearer understanding of what the client requires because it helps us see lines along which a successful design might emerge. That is, we want to *clarify* what the client requires, account for the project stakeholders, and identify contexts within which our design will function. In so doing, we will be *defining* or *framing* the design problem clearly and realistically.

In Chapter 2, we talked about the role of questioning in the design process. A design team may ask questions of the clients and stakeholders who might have varying degrees of interest in the design, including potential users or experts in the field. The experts may be versed in relevant technology or knowledgeable about the market for which the design is aimed. Design teams may also hold their own internal discussions in which they ask each other questions to elicit and list ideas that they can then organize into some problem-relevant structure. It is important that such team discussion sessions remain focused, particularly as they shift from the more general notion of overall design requirements toward the specifics of objectives, constraints, or functions. The best outcome of this work is a list of attributes from which separate lists of *objectives* (i.e., features or behaviors), *constraints* (i.e., limits), and *functions* (i.e., things the design must do) can be extracted.

### 3.2.1 Lists of Design Attributes and of Design Objectives

Imagine that we are on a design team that is consulting for a company that makes both low- and high-quality tools (with a corresponding range of prices). That company's management has given the team a problem statement more specific than the above examples, namely, "Design a new ladder for electricians or other maintenance and construction professionals working on conventional job sites." To fully understand the goals for this design, we have to talk with management, potential users, the company's marketing people, and experts. Building on our previously discussed idea of design as questioning, we ask:

- What features or behaviors would you like the ladder to have?
- What do you want this ladder to do?
- Are there already ladders on the market that have similar features?

And while asking these three questions, we might also ask:

- What do you mean by that?
- How are you going to do that?

**TABLE 3.1 A starting list of safe ladder attributes (O = objectives; C = constraints; F = functions; M = means)**

| Characteristics | O | C | F | M |
|---|---|---|---|---|
| Ladder should be useful | ✓ | | | |
| Used to string conduit and wire in ceilings | ✓ | | | |
| Used to maintain and repair outlets in high places | ✓ | | | |
| Used to replace light bulbs and fixtures | ✓ | | | |
| Used outdoors on level ground | ✓ | | | |
| Used suspended from something in some cases | ✓ | | | |
| Used indoors on floors or other smooth surfaces | ✓ | | | |
| Could be a stepladder or short extension ladder | | | | ✓ |
| A folding ladder might work | | | | ✓ |
| A rope ladder would work, but not all the time | | | | ✓ |
| Should be reasonably stiff and comfortable for users | ✓ | | | |
| Step deflections must be less than 0.05 in. | | ✓ | | |
| Should allow a male of medium height to work safely up to 11-ft. heights | | | | |
| Must support weight of an average worker | | | ✓ | |
| Must be safe | ✓ | | | |
| Must meet OSHA requirements | | ✓ | | |
| Must not conduct electricity | | ✓ | ✓ | |
| Could be made of wood or fiberglass, but not aluminum | | | | ✓ |
| Should be relatively inexpensive | ✓ | | | |
| Should be portable between job sites | ✓ | | | |
| Should be light | ✓ | | | |
| Must be durable | ✓ | | | |

- Why do you want that?
- Are there things or circumstances you want us to avoid?

After our questioning (and perhaps some more team discussions), we might develop a list of attributes for a safe ladder design such as that in Table 3.1. We can categorize the items in our list of attributes into objectives, constraints, functions, and means. It may be useful to review those definitions before we try to categorize the items on our list:

- **objective** *n*: a feature or behavior that the design should have or exhibit.
  *Objectives* are normally expressed as adjectives that capture what the design should *be*, as opposed to what the design should *do*. For example, saying that a ladder should be portable or lightweight expresses an attribute that the client wants the ladder to have. These features and behaviors, expressed in the natural languages of the client and of potential users, make the object "look good" in the eyes of the client or user. We will examine objectives in greater detail in Chapter 4.
- **constraint** *n*: a limit or restriction on the design's behaviors or attributes.
  *Constraints* are clearly defined limits whose satisfaction can be framed into a binary choice (e.g., a ladder material is a conductor or it is not). Any designs that violate these limits are unacceptable. For example, when we say a ladder *must* meet OSHA standards, we are stating a constraint.
- **function** *n*: a specific thing a designed device or system is expected to do.
  *Functions* are typically expressed as "doing" terms in a *verb–noun* pairing. Often they refer to *engineering functions*, such as the second function in Table 3.1: "Must not conduct electricity." Note that this function is also a constraint. We will say much more about functions in Chapter 6.
- **means** *n*: a way or method to make a function happen.
  *Means* or *implementations* are often expressed in very specific terms that, by their nature, are solution-specific. Means often come up because clients or others think of examples of things they've seen that they think are relevant. Because they are so strongly function-dependent, they should be pruned from our attribute list for the time being, but we will revisit them after we have looked at functions.

We also note that some of the entries could be made more precise (and more informative), perhaps by further questioning. For example, the statement "Should allow a male of medium height to work safely up to 11-ft. heights" might well mean "Must allow a 25th percentile male to reach a height of 11 ft." Similarly, "Must support weight of an average worker" might suggest instead "Must support 75th percentile male when on second step from the top." Indeed, as they are now, and even with the possible rewordings, these statements are hard to categorize since they seem to be a mix of objectives, function specifications, and constraints.

We can now separate the list of requirements in Table 3.1 to three separate lists, one of objectives, another of constraints, and a third of functions. We can use these lists for a fuller exploration of each type of attribute. In Chapters 4–6, we will use these lists to help us learn how each type of attribute helps us to be an effective designer.

## 3.3 REVISED PROBLEM STATEMENTS: PUBLIC STATEMENTS OF THE DESIGN PROJECT

We noted early on that design projects are typically initiated with a client's relatively brief statement of what he wants. As we gather information from clients, users, and others, our own views of the problem will shift. In addition to the initial clarification exercise, we will be gathering information that we can present as objectives, constraints, and functions. It is important to recognize the impact of all the new information we've gathered and developed. We can formalize our new (and possibly evolving) understanding by drafting a *revised problem statement* that reflects our fuller understanding of the design problem. A comparison of the initial and revised problem statements for a project often shows a considerable improvement in understanding what the client wants. We will see such results in Section 3.4.

The revised problem statement is also an important tool for developing the team charter that we discuss in Chapter 16 when we consider management of design projects. The team charter is an agreement between the team, their parent institution, and the client regarding what the project is to do, including what is to be delivered. While a charter can be kept fairly general (e.g., with deliverables like "a working prototype"), a thorough revised problem statement allows the team to state more concretely what the client wants and needs.

The revised problem statement is an important communications tool in its own right. The client may have limited technical knowledge about the project topic or the design process. For such a client, the revised problem statement is the "face" of the project: it expresses in clear, unambiguous terms the design problem that the designers are trying to solve. This revised problem statement will often appear in public presentations and reports as well. As such, the revised problem statement deserves serious attention. More than a simple housekeeping task, it presents the client's and the design team's intentions to the world.

A word of caution should be attached to the discussions with the client or the public surrounding the revised problem statement. In our desire to be as accurate as possible, we can be inadvertently insensitive to how our presentations can be perceived by others. Compare the sentence, "After we consulted the client and relevant experts, we adopted and used the following revised problem statement," with "After fixing errors and eliminating our client's bias, we revised the problem statement." The first sentence reflects a collaborative search for the best design solution, while the second one suggests that the client was careless or biased, which is almost certain to embarrass or annoy the client.

## 3.4 DESIGNING AN ARM SUPPORT FOR A CP-AFFLICTED STUDENT

We now begin to describe the design of a device to support and stabilize the arm of a young student afflicted with cerebral palsy (CP) as she writes or draws. We will use this example throughout the book, so it is useful to see how the problem statement originated and was clarified and ultimately revised. The sponsor of this project, the Danbury School, is a special education school within the Claremont (California) Unified School District that serves children with severe orthopedic and medical problems. Students may be as young as three years old; Danbury School offers classes through the sixth grade. Danbury School has a long history of working with the students in Harvey Mudd's introductory design course (E4), dating back to the spring semester in 1992. Among the E4 design projects that have been done for Danbury School are a robotic arm to feed handicapped children, a computer input device for handicapped children, and restrooms for orthopedically disabled students.

In this design problem, E4 teams were asked to design a device for Jessica, a third-grade student who had been diagnosed with CP. (It is worth noting that one of the two teams whose results we cite (team B) used only Jessica's first name in all of their documentation, the other team (A) called her Jane Doe, or Jane for short, to preserve her anonymity.) While a particular user for the design had been identified, Danbury School's principal hoped—as did some of our student designers—that the effort would produce a design that might be developed into a product to be offered to similarly disabled students elsewhere. The full initial problem statement given to the design teams is:

*The Danbury Elementary School of the Claremont Unified School District has a number of students with the diagnosis of CP, a neuro-developmental impairment that causes disturbances of voluntary motor function. For these students, activities that require fine muscle movements (e.g., writing) are particularly difficult because of impaired motor control and coordination as a result of CP. There is ample evidence indicating that these students write more effectively when an instructor physically stabilizes either the hand or the elbow to reduce extraneous movement. A device that can achieve the same physical effect by counteracting the involuntary movement would be desirable since this would increase the students' functional independence.*

A reading of the above initial problem statement makes it clear that the design teams had many questions to answer before they could begin to specify the ultimate form of an arm support. Perhaps the most pressing of these questions is, "what exactly do the client (Danbury School) and the user (Jessica) want (and need)?" To answer this question, the student teams undertook research into cerebral palsy, the personal and classroom environments in which Jessica worked, and into existing designs for arm supports and/or restraints. In addition, teams had to determine what both the client and Jessica meant by phrases such as "more effectively" and "increase . . . independence." The students got their answers through library research, web searches, and interviews with Jessica and Danbury School staff. This problem framing work ultimately led to revised problem statements. Team A produced the following revised problem statement:

*The problem presented to the team involves Jessica, a third-grader at Danbury Elementary School. Jessica has recently begun painting, but because she suffers from cerebral palsy, she has difficulty pursuing her new interest. Jessica paints with her left hand, with her elbow held above the rest position, using a combination of arm and torso movement. While painting, Jessica exhibits exaggerated movements, and lack of control of finer movements, in all directions. These problems are amplified when her arm is fully extended. Currently, when Jessica wants to paint, she requires a teacher or staff member to hold her left elbow stable. The staff at Danbury school has asked the team to try to design a device that would decrease the magnitude of the exaggerations and assist Jessica in controlling her finer movements. The device must permit the same range of voluntary motion currently employed while painting. Thus, the device would take the place of the teacher or staff member and increase Jessica's functional independence while painting in a classroom environment. The Danbury staff must be able to set up the device in a classroom environment in eight minutes or less. Optimally, the device could be used by other students with cerebral palsy or other functionally similar conditions at Danbury Elementary School.*

Design team B produced a revised problem statement with less detail:

*The Danbury Elementary School of the CUSD has a student diagnosed with CP, a neuro-developmental impairment that causes disturbances of voluntary motor function. For this student, activities that require fine muscle movements, such as painting, writing, and eating, are particularly difficult because of impaired motor control and coordination. There is ample evidence indicating that this student paints more effectively when an instructor holds onto the lower portion of the upper arm (right above the elbow) and thus minimizes extraneous movements of the shoulder. The school desires a device that can minimize the student's involuntary shoulder movements and thus allow her to paint semi-independently. Such a device would ideally be applicable in other CP cases and must be easily implemented by an adult.*

These revised problem statements highlight a number of issues surrounding this design exercise and more generally about the practice of engineering design. Notice that team A focused on a specific activity they wish to facilitate (painting), while team B looked at a more general set of activities (painting, writing, and eating). Notice also that team A has identified a constraint (i.e., an 8-minute limit on setup time) that was considered important enough to call out in the revised problem statement. As a result, the client might think these two projects are no longer identical. This shows how we can use the revised problem statement early in the process of problem definition to make sure we are focused on an accurate understanding of the right problem. It also amplifies the open-ended nature of design: two teams working for the same client may come to two different understandings of the same problem. We will see how the design process for both teams A and B unfolds in the next several chapters.

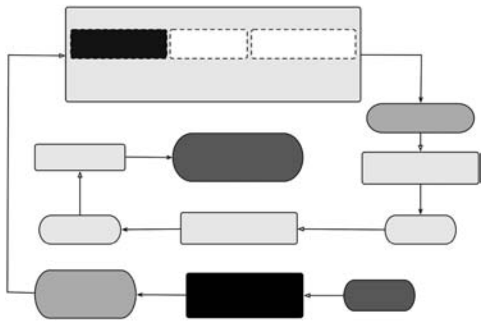
## 3.5 NOTES

*Section 3.4*: The results for the Danbury arm support design project are taken from final reports by Attarian et al. (2007) and Best et al. (2007) submitted during the Spring 2007 offering of Harvey Mudd College's first-year design course, E4: Introduction to Engineering Design. The course is described in greater detail in Dym (1994b).

CHAPTER 4

# PROBLEM DEFINITION: CLARIFYING THE OBJECTIVES

*What is this design intended to achieve?*



IN CHAPTER 3 we focused on the client's initial problem statement, and talked about how we worked to revise that problem statement by (1) identifying and correcting any errors, biases, or implied solutions; and (2) incorporating the appropriate objectives, constraints, and functions. We now turn to clarifying the *objectives* that we want our design to achieve. We will examine the *constraints* that must be satisfied in Chapter 5 and the *functions* the design must perform in Chapter 6.

## 4.1 CLARIFYING A CLIENT'S OBJECTIVES

In Table 3.1 we presented a starting list of the desired attributes of a safe ladder, including objectives, constraints, functions, and means. We have pruned that list of attributes to the shorter list of objectives that we show in Table 4.1, but we see that our list still has a lot of entries. We might find the list more useful if we could organize it in some way. For example, the several uses that we have identified for the ladder might be grouped or clustered together in some coherent way.

Another way to group list entries might be to ask why we care about them. For example, why do we want our ladder to be used outdoors? Maybe that's part of what makes a ladder useful, which relates to another entry on our list. Similarly, we could ask why we care

**TABLE 4.1 A pruned list of objectives for a safe ladder**

| |
|---|
| Ladder should be useful |
| Used to string conduit and wire in ceilings |
| Used to maintain and repair outlets in high places |
| Used to replace light bulbs and fixtures |
| Used outdoors on level ground |
| Used suspended from something in some cases |
| Used indoors on floors or other smooth surfaces |
| Should be reasonably stiff and comfortable for users |
| Should allow a person of medium height to reach and work at levels up to 11 ft. |
| Must be safe |
| Should be relatively inexpensive |
| Must be portable between job sites |
| Should be light |
| Must be durable |

whether the ladder is useful. In this case, the answer is not on the list: We want it to be useful so that people will buy it. Put another way, usefulness makes a ladder marketable. This suggests that we need an entry on marketing for our pruned objectives list: "The ladder should be marketable." This turns out to be a very helpful objective, since it tells us why we want the ladder to be cheap, portable, etc. (On the other hand, we should be careful identifying "superobjectives" such as marketability, since almost any new or interesting product feature could fit under that rubric.) If we go through *thoughtful* clustering of our questions in this way, we can develop a new list that we can represent in an *indented outline*, with *hierarchies* of major headings and various levels of subheadings (e.g., Table 4.2).

**TABLE 4.2 An indented list of the pruned objectives for a safe ladder**

**0. *A safe ladder for electricians***
- **1. The ladder should be safe**
  - 1.1 The ladder should be stable
    - 1.1.1 Stable on floors and smooth surfaces
    - 1.1.2 Stable on relatively level ground
  - 1.2 The ladder should be reasonably stiff
- **2. The ladder should be marketable**
  - 2.1 The ladder should be useful
    - 2.2.1 The ladder should be useful indoors
      - 2.2.1.1 Useful to do electrical work
      - 2.2.1.2 Useful to do maintenance work
    - 2.2.2 The ladder should be useful outdoors
    - 2.2.3 The ladder should be of the right height
  - 2.2 The ladder should be relatively inexpensive
  - 2.3 The ladder should be portable
    - 2.3.1 The ladder should be light in weight
    - 2.3.2 The ladder should be small when ready for transport
  - 2.4 The ladder should be durable

The revised, indented outline in Table 4.2 allows us to explore each of the higher-level objectives further, in terms of the subobjectives that tell us how to realize them. At the top level, our objectives turn us back to the original design statement we were given, namely to design a safe ladder that can be marketed to a particular group.

We have certainly not exhausted all of the questions we could ask about the ladder, but this outline answers some of the questions mentioned earlier. For example, "What do you mean by safe?" is answered by two subobjectives in the cluster of safety issues: The designed ladder should be both stable and relatively stiff. We have answered "How are you going to do that?" by identifying several subobjectives or ways in which the ladder could be useful within the "The ladder should be useful" cluster and by specifying two further "sub-subobjectives" about how the ladder would be useful indoors. And we have answered the question "Why do you want that?" by indicating that the ladder ought to be cheap and portable in order to reach its intended market of electricians and construction and maintenance specialists.

### 4.1.1 Representing Lists of Objectives in Objectives Trees

The indented outline of objectives in Table 4.2 is one way to represent the information contained in that list. That same information can also be represented or portrayed graphically in a *hierarchy* of boxes, each of which contains an objective for the object being designed, as shown in Figure 4.1. Each layer or row of objective boxes corresponds to a level of indentation (indicated by the number of digits to the right of the first decimal point) in the outline. Thus, the indented outline becomes an *objectives tree*: a graphical

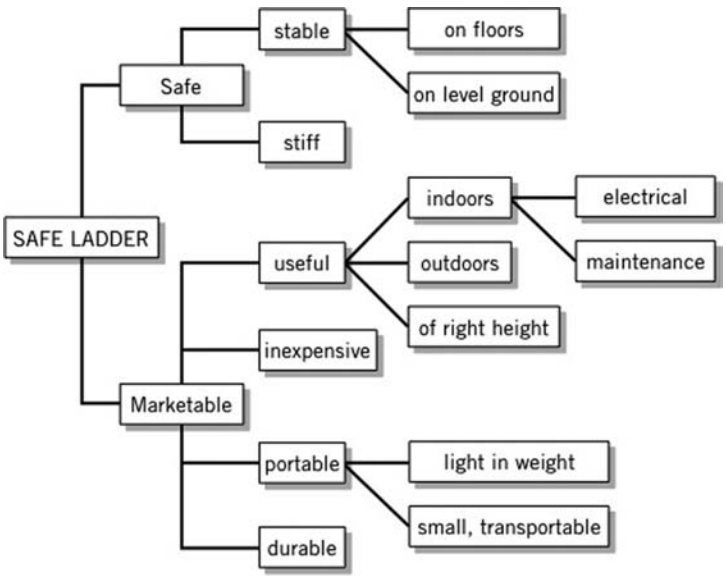



**Figure 4.1** The objectives tree for the design of a safe ladder. Note the hierarchical structure and the clustering of similar ideas.

depiction of the *objectives for the device or system*. The top-level goal in an objectives tree—the *root* node at the top of the tree—is *decomposed* or broken down into subobjectives at differing levels of importance or to include progressively more detail. Thus, the tree reflects a *hierarchical structure* as it expands downward. An objectives tree also gives the tree some organizational strength and utility by *clustering* together related subobjectives or similar ideas.

The graphical tree display is very useful for portraying design issues, and for highlighting things we need to measure, since these objectives will provide our basis for choosing between alternatives. The tree format also corresponds to the mechanics of the process that many designers follow: One of the most useful ways of "getting your mind around" a large list of objectives is to put them all on Post-It$^{TM}$ notes, and then move them around until the tree makes sense. Note, too, that process just outlined—from lists to refined lists to indented outlines to trees—has a lot in common with outlining, a fundamental skill of writing. A topical outline provides an indented list of topics to be covered, together with the details of the subtopics corresponding to each topic. Since each topic represents a goal for the material to be covered, the identification of an objectives tree with a topical (or an indented) outline seems logical.

### 4.1.2 Remarks on Objectives Trees

In addition to their use in depicting design objectives, objectives trees are valuable in several other ways. First, and perhaps foremost, note that as we work *down* an objectives tree (or further in on the levels of indentation of an outline), we are not only getting more detail. We are also answering a generic *how* question for many aspects of the design: "*How* are you going to do that?"

Conversely, as we move *up* the tree, or further out toward fewer indentations, we are answering a generic *why* question about a specific objective: "*Why* do you want that?" This may be important if, when selecting a design, we find that one alternative is better with respect to one objective, but weaker with respect to another.

But if we're working downward as we construct and organize a tree, where do we stop? When do we end our list or tree of objectives? One simple answer is: We stop when we run out of objectives and implementations begin to appear. That is, within any given cluster, we could continue to parse or decompose our subobjectives until we are unable to express succeeding levels as further subobjectives. The argument for this approach is that it points the objectives tree toward a *solution-independent* statement of the design problem. We know what characteristics the design has to exhibit, without having to make any judgment about how it might get to be that way. In other words, we determine the features or behaviors of the designed object without specifying the way the objective is realized in concrete form.

We can also limit the depth of an objectives tree by watching for verbs or "doing" words because they normally suggest functions. Functions do not generally appear on objectives trees or lists.

Another tree-building issue has to do with deciding what to do with the things that we have removed from the original list of attributes. The functions and means (or implementations) are simply put aside—recorded, but not discarded—to be picked up again later in the process. Constraints, however, are sometimes added to objectives tree, although in ways that very clearly distinguish them from the objectives: We might present constraints in boxes differently shaped than the objectives. In an outline form of the objectives tree, we might use italics or a different font to denote constraints (see Figure 5.1 in Section 5.2). In either case, it is most important to recognize that constraints are related to but are different from objectives: They mean very different things and are used in different ways.

Obviously, it is important to take notes when we are generating our lists of objectives, because we are generating a lot of information, to ensure that *all* suggestions and ideas are captured, even those that seem silly or irrelevant at the moment. Then it becomes important to organize the information we're getting so we can use it effectively: It's always easier to prune and throw away things than to recapture spontaneous ideas and inspirations. Also, get the substance of the objectives down first: Once a rough outline of an objectives tree has emerged, it can be formalized and made to look presentable and pretty with any number of standard software packages for constructing organization charts or similar graphical displays.

Finally, do we build an objectives tree as soon as we start a design job, or after doing some homework and learning more about the design task we're undertaking? There's no hard and fast answer to these questions, in part because building an objectives list or tree is not a mathematical problem with an attendant set of initial conditions that must be met. Also, building a tree is not a one-time, let's-get-it-done kind of activity. It's an iterative process, but one that a design team should start with at least some degree of understanding of the design domain. Thus, some of the questioning of clients, users, and experts should have begun, and some of the tree building can go on episodically while more information is being gathered.

### 4.1.3 The Objectives Tree for the Juice Container Design

In the juice container design example, our design team is working for one of the two competing food product manufacturers.

In questioning people such as the marketing staff and potential customers to clarify what was wanted from this design, we might have heard several motivations driving the desire for a new "juice bottle," including:

- plastic bottles and containers all look alike;
- the product has to be delivered to diverse climates and environments;
- safety is a big issue for parents whose children might drink the juice;
- customers, especially parents, are concerned about environmental issues;
- the market is very competitive;
- parents (and teachers) want children to be able to get their own drinks; and
- children always spill drinks.

This list of motivations emerged during the questioning process, and their effects are displayed in an *annotated attributes list* for the container in Table 4.3. This annotated list also shows how that some of the listed objectives are expanded into subobjectives, while others are connected to existing objectives at higher levels. In one case a brand new top-level objective, Highly Marketable, is identified. We show in Figure 4.2 the objectives tree corresponding to (and expanded from) the annotated attribute list of Table 4.3. The detailed

**TABLE 4.3 An annotated objectives list for the juice container design**

| | | |
|---|---|---|
| Safe | ⟶ | DIRECTLY IMPORTANT |
| Perceived as Safe | ⟶ | Appeals to Parents |
| Inexpensive to Produce | ⟶ | Low Production Cost |
| Low Production Cost | ⟶ | DIRECTLY IMPORTANT |
| Distinctive Appearance | ⟶ | Strong Brand Identity |
| Environmentally Benign | ⟶ | Safe |
| Environmentally Benign | ⟶ | Appeals to Parents |
| Long Shelf Life | ⟶ | Highly Marketable |
| Easy for Kids to Use | ⟶ | Appeals to Parents |
| Temperature Resistant | ⟶ | Durable for Shipment |
| Shock Resistant | ⟶ | Durable for Shipment |
| Easy to Distribute | ⟶ | Low Production Cost |
| Durable for Shipment | ⟶ | Easy to Distribute |
| Easy to Open | ⟶ | Easy for Kids to Use |
| Hard to Spill | ⟶ | Easy for Kids to Use |
| Appeals to Parents | ⟶ | Highly Marketable |
| Strong Brand Identity | ⟶ | Highly Marketable |
| Highly Marketable | ⟶ | DIRECTLY IMPORTANT |

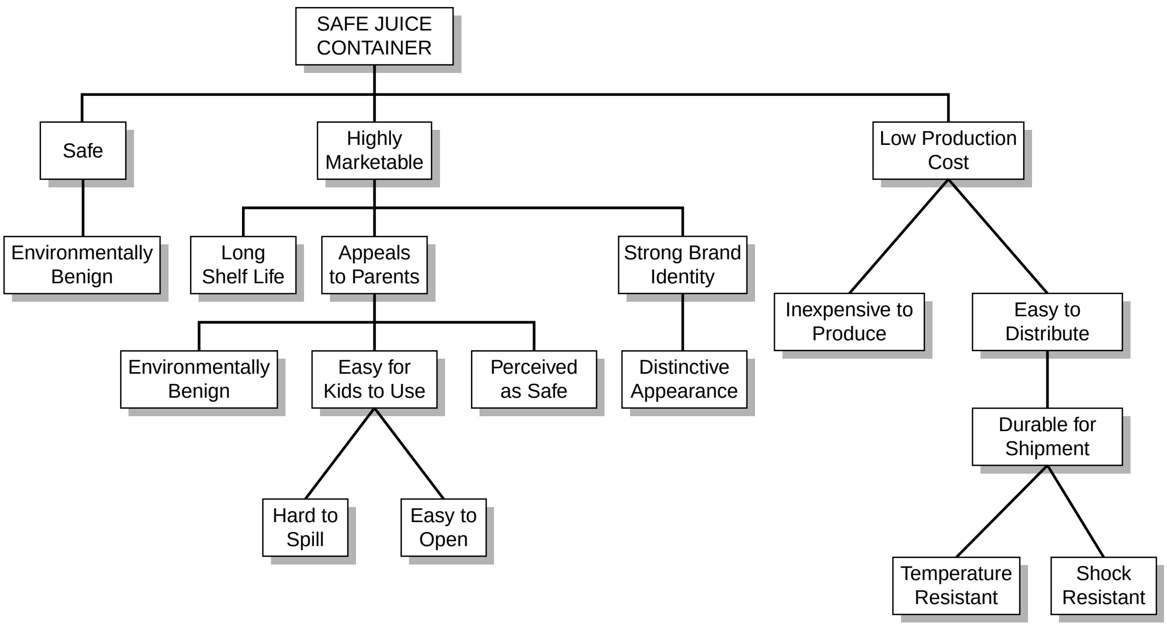



**Figure 4.2** An objectives tree for the design of a new juice container showing the hierarchical structuring of the needs identified by the juice company and by the potential consumers of the new juice drink.

subobjectives that emerge in these trees clearly track well with the list of concerns and motivations identified in the clarification process.

We also note in Figure 4.2 that some of the very detailed subobjectives are written as passive, "are" descriptions (e.g., Shock Resistant, Temperature Resistant). These subobjectives might easily be written in more active forms, for example, Resist Shock and Resist Temperature. This confirms our statement in Section 4.1.2 that we get to *how* we achieve objectives as we move down an objectives tree, and also anticipates how we express functions (see Section 6.2.5).

## 4.2 MEASUREMENT ISSUES IN ORDERING AND EVALUATING OBJECTIVES

Having identified the client's objectives for a design, we ask: Are some objectives more important than others, that is, what are the client's priorities? A second question arises, how will we know whether objectives have been achieved? These two questions imply the third question: Are there measurements we could make to compare design objectives and their relative achievement? This last question will be central when we near the end of the conceptual design phase and choose a single, final design, so we defer discussion of the third question until Chapter 8. We will answer the first two in Sections 4.3 and 4.4, respectively, after first providing some needed context about measurement.

Engineers are used to measuring all sorts of things: beam lengths, surface areas, hole diameters, speeds, temperatures, pressures, and so on. In each of these cases, there is a ruler or scale involved that shows a zero and has marks that show units, whether they be inches, microns, millimeters of mercury, or degrees Fahrenheit or centigrade. The ruler establishes a common basis for comparison. Without rulers, we cannot meaningfully quantify the assertion that "*A* is taller than *B*" short of simply standing *A* and *B* against each other, back to back. However, by using a measuring stick that is marked with a zero and a countable number of intervals of fixed length, we can establish real numbers to represent their heights and make meaningful comparisons of their heights.

The important point here is that of having a *ruler* or *scale* with (1) a *defined zero*, and (2) a *unit* to define the markings scribed onto the ruler. In mathematical terms, these properties enable *strong measurements* that allow us to treat measured mathematical variables (say $L$ for length, $T$ for temperature, and so on) as we would any variable in calculus. Thus, strong measurements can be used like any of our "normal" physical variables in a mathematical model.

On the other hand, we generally use *ordinal scales* to place things in rank order, that is, in first, second, or $n$th place. This is not as straightforward or useful as it seems. Suppose we're standing by a race's finish line, but without a watch: We can tell the order of the racers' arrivals, but not their respective times. Thus, we can say that *A* is faster than *B*, who is faster than *C*, and so on—but we can never say by how much. Similarly, people say that they prefer vanilla ice cream to chocolate, *but not by how much*. This is precisely the problem with asking the client to set priorities because we are really asking for a *subjective ranking* of their relative importance. The client may prefer portability over price in designing a new laptop computer, but there is no meaningful way to say that "portability is five times more important than cost" because there is no scale or ruler that defines both a zero and a unit with which to make such measurements.

## 4.3 RANK ORDERING OBJECTIVES WITH PAIRWISE COMPARISON CHARTS

We have been rather insistent in this chapter that we properly identify and list all of the client's objectives, while taking great care to separate out constraints, functions, or means. But do we know that all of the identified objectives have the same importance or value to the client or the users? Since we made no effort to see whether there is any variation in an objective's relative value, we have implicitly assumed that every top-level objective has the same value to all concerned. Yet it is almost certain that some objectives are more important than others, so we ought to be able to recognize that and measure it. How are we going to do that?

### 4.3.1 An Individual's Rank Orderings

Suppose we want to know the relative value or importance of objectives, one to another, and then order them accordingly. Sometimes the client will state clear preferences, or perhaps a potential user does, so that we, as designers, don't have to determine an explicit ordering ourselves. More often, however, we have to elicit values from the client. Fortunately, there is a straightforward technique that can be used by an individual to *rank* objectives that are on the same level in our objectives tree or are within the same grouping or cluster. It is very important that we make our comparisons of objectives with this hierarchical restriction firmly in mind to ensure that we're comparing apples with apples and oranges with oranges. For example, rank ordering the ladder's usefulness, cost, portability, and durability would provide useful design information. On the other hand, it makes little sense to compare having a ladder be useful for electrical work against its durability.

Consider a ladder with the four high-level objectives stated cost, useful, portable, and durable. The *pairwise comparison chart* (PCC) is a tool for ordering the relative importance of objectives. It is based on the assumption that we can order any two objectives taken as a pair. For example, we prefer cost to durability, portability to cost, portability to convenience, and so on. The PCC is a simple matrix that allows us to (1) compare every objective with each remaining objective individually, and (2) add total scores for each objective.

Table 4.4 shows a PCC for our four-objective ladder design. The entries in each box of the chart are determined as binary choices: Every entry is either a 1 or a 0, where 1 indicates that the row objective is preferred over the column objective. Along the row of each objective, such as cost, we enter a zero in the columns for portability and convenience if they are preferred over cost, and we enter 1 in the durability column because cost is preferred over durability. We enter nothing in the diagonal boxes corresponding to

**TABLE 4.4 A pairwise comparison chart (PCC) for a ladder design**

| Goals | Cost | Portability | Convenience | Durability | Score |
|---|---|---|---|---|---|
| Cost | •••• | 0 | 0 | 1 | 1 |
| Portability | 1 | •••• | 1 | 1 | 3 |
| Convenience | 1 | 0 | •••• | 1 | 2 |
| Durability | 0 | 0 | 0 | •••• | 0 |

weighting any objective against itself, and we enter ratings of 0.5 for objectives valued equally. The scores for each objective are found simply by adding across each row. Here the four objectives are ranked (with their scores) in order of decreasing value or importance: portability (3), convenience (2), cost (1), and durability (0).

Note, too, that the score of 0 earned by durability does *not* mean that we can or should drop it as an objective! Durability earned the 0 because it was ranked as *least important*, that is, it placed last in the line of the four objectives ranked. If it were of *no importance*, it would not have been listed as an objective to begin with. Thus, we *cannot* drop objectives that score zeroes.

It should also be noted that the pairwise comparison, if done correctly, preserves the important property of *transitivity*. That is, in the ladder design we preferred portability to convenience and convenience to cost—and the PCC produced a consistent result when it said that we preferred portability to cost. If we have inconsistencies in our PCC (e.g., if we preferred cost to portability), we should clear up exactly what the client's preferences are. The simple PCC process just described, which is also known as the Borda count, is a valid way of ordering things, but its results should be taken as *no more than a straightforward rank ordering*, or an ordering of place in line. The scores assembled in Table 4.4 do not constitute what we had defined as strong measurement because there is no scale on which we can measure the four objectives, and the zero is only implied, not defined. We cannot, however, say that we portability is three times as important as cost or twice as important as convenience.

## 4.3.2 Aggregating Rank Orderings for a Group

Sometimes we need to develop an aggregate ranking for a group of clients, users, or designers. We have so far worked in the framework of the single decision maker who is making a subjective assessment to obtain a meaningful and useful ranking. The group situation—in which a group of clients or users (or a design team) collects individual votes to aggregate into a set of preferences for the entire group (or team)—is still more complicated and a subject of both research and discussion. The sticking point derives from the *Arrow Impossibility Theorem* of decision theory, for which Kenneth J. Arrow won the Nobel Prize in Economics in 1972. It states, in essence, that it is impossible to run a "fair" aggregation and preserve transitivity when there are more than two objectives to rank. There is corresponding discussion in the design community as to the role that decision theory plays in the design process, but we believe that the PCC (or Borda count) can be used to indicate the collective preferences of a group of clients or of a design team.

Suppose a team of 12 people is asked to rank order three objectives: $A$, $B$, and $C$. In doing so, the 12 individuals produce 12 individual orderings that, using the ranking symbol $\succ$ to indicate that $A \succ B$ means "$A$ is preferred to $B$," are

$$\begin{array}{ll} 1 \text{ preferred } A \succ B \succ C & 4 \text{ preferred } B \succ C \succ A \\ 4 \text{ preferred } A \succ C \succ B & 3 \text{ preferred } C \succ B \succ A \end{array} \tag{4.1}$$

The collective will of these 12 individuals is worked out through the aggregated PCC in Table 4.5. One point is awarded to the winner of each pairwise comparison, and the number

**TABLE 4.5 An aggregated pairwise comparison chart (PCC) for 12 people**

| Win/Lose | *A* | *B* | *C* | Sum/Win |
|---|---|---|---|---|
| *A* | •••• | 1 + 4 + 0 + 0 | 1 + 4 + 0 + 0 | 10 |
| *B* | 0 + 0 + 4 + 3 | •••• | 1 + 0 + 4 + 0 | 12 |
| *C* | 0 + 0 + 4 + 3 | 0 + 4 + 0 + 3 | •••• | 14 |
| Sum/Lose | 14 | 12 | 10 | •••• |

of points awarded to each alternative by each of the 12 rankers is summed. The aggregate rank ordering of the three objectives is

$$C \succ B \succ A \tag{4.2}$$

The group consensus, based on summing, is that *C* is most important, *B* second, and *A* least. The 12 choose objective *C* as their most important objective, even though it clearly was not unanimous. In fact, only 3 of 12 designers ranked it most important. However, the PCC as applied here provides as good a tool as there is for these purposes, so long as its results are used with the same caution noted for individual PCCs.

### 4.3.3 Using Pairwise Comparisons Properly

The role of PCCs in design is still a subject of design research, in part, because it has often been misstated and/or misapplied. Thus, we want to repeat and add some cautionary notes to ensure their proper use. First, the PCC approach should be applied in a *constrained*, *top-down* fashion, so that (1) objectives are compared only when at the same level on the objectives tree, and (2) the higher-level objectives are compared and ranked before those at lower, more detailed levels. The second point seems only a matter of common sense to ensure that more "global" objectives (i.e., those more abstract objectives that are higher up on the objectives tree) are properly understood and ranked before we fine-tune the details. Also, for many design tasks, only top-level objectives need be so ranked. It would make sense to rank objectives below the top level only for the design of complex subsystems, within large and complex systems.

In addition, given the subjective nature of such rankings, we should ask whose values are being assessed when we use a PCC. For example, marketing values can easily be included: A ladder design team might want to know whether it's "better" for a ladder to be cheaper or heavier. On the other hand, there could be objectives rankings that reflect fundamental values of clients and/or designers. For example, consider how the objectives for the juice container design might be ranked at two competing companies, ABC and NBC. We show PCCs for the ABC- and NBC-based design teams in parts (a) and (b) of Table 4.6, respectively. These two charts and the scores in their right-hand columns, show that ABC was far more interested in a container that would generate a strong brand identity and be easy to distribute than in one that would be environmentally benign or appeals to parents. At NBC, on the other hand, the environment and the taste preservation ranked more highly. Thus, subjective values show up in PCCs and, consequently, in the marketplace!

Finally, one more cautionary warning: Rankings of objectives *cannot* be put on a scale or ruler. We *cannot* attach relative weights to objectives or make similar calculations. We can't

**TABLE 4.6 PCCs for the design of a new juice container ranking objectives at (and reflecting the values of) (a) ABC and (b) NBC**

| Goals | Environ. Benign | Easy to Distribute | Preserve Taste | Appeals to Parents | Market Flexibility | Brand ID | Score |
|---|---|---|---|---|---|---|---|
| Environ. Benign | •••• | 0 | 0 | 0 | 0 | 0 | 0 |
| Easy to Distribute | 1 | •••• | 1 | 1 | 1 | 0 | 4 |
| Preserve Taste | 1 | 0 | •••• | 0 | 0 | 0 | 1 |
| Appeals to Parents | 1 | 0 | 1 | •••• | 0 | 0 | 2 |
| Market Flexibility | 1 | 0 | 1 | 1 | •••• | 0 | 3 |
| Brand ID | 1 | 1 | 1 | 1 | 1 | •••• | 5 |

(a)

| Goals | Environ. Benign | Easy to Distribute | Preserve Taste | Appeals to Parents | Market Flexibility | Brand ID | Score |
|---|---|---|---|---|---|---|---|
| Environ. Benign | •••• | 1 | 1 | 1 | 1 | 1 | 5 |
| Easy to Distribute | 0 | •••• | 0 | 0 | 1 | 0 | 1 |
| Preserve Taste | 0 | 1 | •••• | 1 | 1 | 1 | 4 |
| Appeals to Parents | 0 | 1 | 0 | •••• | 1 | 1 | 3 |
| Market Flexibility | 0 | 0 | 0 | 0 | •••• | 0 | 0 |
| Brand ID | 0 | 1 | 0 | 0 | 1 | •••• | 2 |

(b)

answer questions such as "*How much more* important is portability than cost in our ladder?" While there are cases where one objective is far more important than any of the others (e.g., safety for an air traffic control system), there is no mathematical foundation for scaling or normalizing the rankings obtained with tools such as the PCC. The numbers obtained with a PCC are *subjective* orderings of relative value. Therefore, we should not try to make these numbers seem more important by doing further calculations with them (e.g., adding weights) or by giving them unwarranted precision. In fact, to weight objectives is to commit a very clear error of building an appealing numerical edifice on a mathematically unsound foundation.

## 4.4 DEVELOPING METRICS TO MEASURE THE ACHIEVEMENT OF OBJECTIVES

Having determined what our client wants in a design in terms of rank-ordered objectives, we now take up the question of assessing how well a particular design *actually does* all these things. As we noted in Section 4.2, such assessment requires *metrics*, standards that measure the extent to which a design's objectives are realized. In principle, it is easy to devise metrics, as we need only units and a scale of something that can be *measured* about an objective and a way to *assign* a value to the design in terms of those units. In practice, it is often hard to devise and apply an appropriate metric.

Sometimes metrics are straightforward. If we want to minimize the number of parts, we simply count the parts in a design. If we want to minimize manufacturing costs, we estimate its manufacturing costs in the currency of interest. Problems arise, however, in two different ways. First, even when we can actually calculate or measure the achievement of each of our objectives, how do we put their different metrics (e.g., part counts, manufacturing dollars) on a common scale so that we can compare their respective achievements? Second, what do we do in situations where there is no ruler? For example, how would we measure the achievement of "simplicity" as a design objective? We will introduce *value scales* to answer the first question, and *surrogate metrics* to answer the second.

### 4.4.1 Establishing Good Metrics for Objectives

First and foremost, a *metric* should *actually measure the objective* that the design is supposed to meet. Often designers try to measure something that may be interesting, but that is not really on point for the desired objective: Measuring the number of colors on a package may not be a good metric by which to assess its appeal to consumers. On the other hand, sometimes we invoke *surrogate metrics* because there are no obvious measures appropriate to the objective of interest. For example, we might drop a cell phone from several different heights and check its post-drop performance to assess its durability. Similarly, the simplicity (or the complexity) of a product might be assessed in terms of the number of parts that are needed to make the product, or perhaps in terms of the product's estimated assembly time. Thus, surrogate metrics are quite useful when they are measurable properties that strongly relate to the objective of interest.

Having decided *what* to measure, we then determine the *appropriate units* in which to make the measurement(s). For an objective of low ladder weight, for example, we could use units of weight or mass, that is, kg, lb, or oz. For an objective of low cost, our metric would be measured in currency, that is, US$ in the United States. Along with determining appropriate units, we must also be sure that the metric enables the *correct scale* or *level of precision*. For a low-weight ladder, we should not measure weight in either tons or milligrams.

Our next step is to *assign points* for the metric that corresponds to the scale or range expressed in the right *units of interest* or *figures of merit*. For example, if we want a fast car, we might use speed in kilometers per hour as the figure of merit and assume that the range of speed of interest is from 50 to 200 km/h. Then we could assign points linearly distributed over the range, that is, from 0 points at the low end (50 km/h), up to 10 points at the high end (200 km/h). Thus, a design alternative that has a projected speed of 170 km/h would earn or be awarded eight points. Were we assessing the durability of a cell phone, we might drop the phone over a range of heights of $1\,\text{m} \leq \text{height} \leq 10\,\text{m}$ and then assign points from 0 points at the low end (1 m) up to 10 points at the high end (10 m).

Note that in awarding points for speeds or drop test heights, we are implicitly assuming that we have a plan for measuring performance that is compatible with the type of scale and units selected. Such a measurement plan could include laboratory tests, field trials, consumer responses to surveys, focus groups, and so on. But while some things are relatively easy to measure directly (e.g., weight, on a balance scale) or indirectly (e.g., weight, by

**TABLE 4.7 Scales or rulers for awarding points depending on perceived value of a solution (*Use-Value Analysis*) or perceived value of the idea or concept (*VDI 2225* Guidelines)**

| Use-Value Analysis | | VDI 2225 Guidelines | |
|---|---|---|---|
| Solution Value | Points Awarded | Perceived Value | Points Awarded |
| Absolutely useless | 0 | Unsatisfactory | 0 |
| Very inadequate | 1 | | |
| Weak | 2 | Just tolerable | 1 |
| Tolerable | 3 | | |
| Adequate | 4 | Adequate | 2 |
| Satisfactory | 5 | | |
| Good, w/ drawbacks | 6 | Good | 3 |
| Good | 7 | | |
| Very good | 8 | | |
| Exceeds requirements | 9 | Very good (ideal) | 4 |
| Excellent | 10 | | |

computing volume), others must be *estimated* (e.g., the top speed of a planned airplane, by doing a back-of-the-envelope calculation). Further, some things are neither easily measured nor easily estimated (e.g., cost may be hard to estimate if we don't know how we're manufacturing something or how many units we're making, and so on).

Often the appropriate "units" are general categories (e.g., "high," "medium," or "low") or qualitative characterizations (e.g., "great," "okay," or "lousy"). In Table 4.7 we show two ways to quantify qualitative characterizations by assigning "measurement" points to the values or categories. Eleven categories of the value of a solution are offered in *Use-Value Analysis*, with points then being awarded on a scale ranging from 0 (absolutely useless) to 10 (ideal). There are five categories in the German *VDI 2225* standard, with points awarded on a scale ranging from 0 (unsatisfactory) to 4 (very good/ideal), depending on the degree to which an idea or a concept or something else is considered valuable.

Consider once again the ladder objective of low cost. We may not be able to get the information needed to accurately calculate the ladder's manufacturing costs without undertaking a significant, expensive study. We might instead estimate the manufacturing costs by summing the costs of the ladder's components when purchased in given lot sizes. This may neglect some relevant costs (e.g., component assembly, company overhead), but it does allow us to distinguish between designs with expensive elements and designs with cheaper elements. Alternatively, we might seek expert input from our client and then rank the designs into qualitative categories such as "very expensive," "expensive," "moderately expensive," "inexpensive," and "very cheap."

It is important that we measure the achievements of all of the objectives of a design consistently, on the same ruler or scale: We can't allow some objectives to dominate the

**TABLE 4.8 Measuring quantitative performance levels for figures of merit of mass per unit power (kg/kW) and for service life (km) measured on the *Use-Value Analysis* and *VDI 2225* scales or rulers**

<table>
<tr><th colspan="2">Measured/Estimated Values</th><th colspan="2">Value Scaled</th></tr>
<tr><th>Mass/Power (kg/kW)</th><th>Service Life (km)</th><th>Use-Value Points</th><th>VDI 2225 Points</th></tr>
<tr><td>3.5</td><td>$20 \times 10^3$</td><td>0</td><td rowspan="2">0</td></tr>
<tr><td>3.3</td><td>$30 \times 10^3$</td><td>1</td></tr>
<tr><td>3.1</td><td>$40 \times 10^3$</td><td>2</td><td rowspan="2">1</td></tr>
<tr><td>2.9</td><td>$60 \times 10^3$</td><td>3</td></tr>
<tr><td>2.7</td><td>$80 \times 10^3$</td><td>4</td><td rowspan="2">2</td></tr>
<tr><td>2.5</td><td>$100 \times 10^3$</td><td>5</td></tr>
<tr><td>2.3</td><td>$120 \times 10^3$</td><td>6</td><td rowspan="2">3</td></tr>
<tr><td>2.1</td><td>$140 \times 10^3$</td><td>7</td></tr>
<tr><td>1.9</td><td>$200 \times 10^3$</td><td>8</td><td rowspan="3">4</td></tr>
<tr><td>1.7</td><td>$300 \times 10^3$</td><td>9</td></tr>
<tr><td>1.5</td><td>$500 \times 10^3$</td><td>10</td></tr>
</table>

overall assessments by virtue of their being measured on scales that award more points than could be earned by other objectives. In fact, we can use the *Use-Value Analysis* and *VDI 2225* Guidelines of Table 4.7 to ensure that we are assessing quantitative performance ratings on similar, consistent scales. In Table 4.8 we show how two different sets of quantitative performance ratings in their figures of merit, for mass per unit power (measured in kilogram/kilowatt) and service life (measured in kilometer), each of which is arrayed against the *Use-Value Analysis* and German *VDI 2225* scales.

It is important to determine whether or not the information derived from using a metric is worth the cost of actually performing a measurement. The value of the metric may be small in comparison with the resources needed to obtain the measurement. In those cases we can either develop a new metric or find another means for measuring the expensive metric, or look for an alternative way to assess our design. There may be other metrics that provide equivalent information, in which case we may be able to choose a less expensive measurement. In other cases, we may decide to use a less accurate method to assess our designs. As a last resort, we may even decide to convert a hard-to-measure objective into a constraint, which allows us to consider some designs and reject others. (And we must keep in mind the distinction between *converting* objectives to constraints and *confusing* objectives with constraints.) In the case of designing a low-cost ladder without adequate cost information, perhaps that objective could be converted into a constraint such as “Contains no parts costing greater than $20.” This constraint indirectly works toward the original objective, while enabling the dismissal of designs that seem certain to not be low in cost.

A few closing comments about metrics. First, a metric should be *repeatable*: Anyone conducting the same measurement should get the same results, subject to normal experimental error. Repeatability can be fostered either by using standard methods and instruments, or, if they're unavailable, by carefully documenting the protocols followed. We should also run as many tests as we need to ensure the statistical validity of our measurements. Second, we should express our test results in *understandable units of measure*. And, finally, we should work very hard to ensure the *unambiguous interpretation* of our results: Everyone viewing results should come to the same conclusion about the measurement. We certainly don't want a posttest debate about what our assessment or our measurements mean.

### 4.4.2 Establishing Metrics for the Juice Container

Let us now establish metrics for the six objectives (identified in Figure 4.2) for the juice container design. We see immediately that these objectives will require *qualitative* metrics simply because there is no direct measurement that we can make for any of them. Thus, we will establish metrics analogous to *Use-Value Analysis* and the *VDI 2225* Guidelines. We will also keep in mind what we've said before: Metrics should be solution-independent, that is, they should be established without reference to specific design solutions.

Consider first that we want the juice container to be *environmentally benign*. Products that are environmentally benign must at worst do no harm to the environment, that is, they should produce no hazardous waste or residue. At best, containers should be easily reused, or—and almost as good—their materials should be recyclable. Thus, we might propose the following qualitative metric:

Objective: *Juice container should be environmentally benign*.

Units: Rating assessment of most environmentally desirable alternative from 0 (worst) to 100 (best).

Metric: Assign points according to the following scale:

| | |
|---|---|
| Completely reusable: | 100 points |
| Material is recyclable: | 90 points |
| Material is easily disposable: | 50 points |
| Material is disposable with difficulty: | 25 points |
| Material is a hazardous waste: | 0 points |

We might also establish a surrogate metric of *environmental costs* for this objective by, for example, ascertaining the cost of washing bottles to enable reuse (as they once were for sodas and other drinks). Similarly, we might estimate the cost of recycling the materials (i.e., glass bottles and aluminum cans). Finally, the social and opportunity costs of disposal might be estimated. For example, we might consider the cost of disposing of relatively benign materials (e.g., cardboard), or more hazardous materials and products (e.g., the plastic bags that are a major pollutant of the oceans, or small seed-like detritus that is eaten by unsuspecting birds). Then a quantitative surrogate metric can be established using known or estimated environmental costs. Note that environmental, life cycle, and sustainability issues are increasingly central in product design, as we will discuss in Chapter 14.

We also want the juice container to be *easy to distribute*. Thus, we should consider whether the container is easily *packed*, in terms of both shape and size; whether it is *breakable*; and whether it or the juice product is sensitive to *temperature*. It is also likely that standard container shapes make it easier for storeowners to provide shelf space for the new juice. We propose the following use-value analog:

Objective: *Juice container should be easy to distribute*.

Units: Rating of design team's assessment of the ease of packing and stacking the container, from 0 (worst) to 100 (best).

Metric: Assign points according to the following scale:

| | |
|---|---|
| Very easy to pack and stack: | 100 points |
| Easy to pack and stack: | 75 points |
| Can be packed and stacked: | 50 points |
| Hard to pack and stack: | 25 points |
| Very hard to pack and stack: | 0 points |

This metric is one for which a juice company would likely have both data and experience with what works and what doesn't. Clearly, a data-driven metric is much more meaningful—and persuasive—than the qualitative metric we are proposing here.

We also want our juice container to *have a long shelf life*. For this objective, the design team can use:

Objective: *Juice container should have long shelf life*.

Units: Rating of how long the juice product remains in acceptable condition, from 0 (worst) to 100 (best).

Metric: Assign points according to the following scale:

| | |
|---|---|
| Shelf life 1 year (12 months): | 100 points |
| Shelf life 9 months: | 75 points |
| Shelf life 6 months: | 50 points |
| Shelf life 3 months: | 25 points |
| Shelf life 1 month: | 0 points |

This is another instance in which the juice company is almost certain to have both experience and data. In fact, for this objective and the remaining three objectives (*appeal to parents*, *permit marketing flexibility*, and *generate brand identity*), our design team is almost certainly going to ask our client's marketing teams and other in-house resources (e.g., ABC and NBC senior managers) for the information we need to assess the attainment of these objectives. This may also be an occasion to use some of the established ways (described in Section 2.3.3) of assessing the market (e.g., focus groups, structured questionnaires, surveys).

## 4.5 OBJECTIVES AND METRICS FOR THE DANBURY ARM SUPPORT

The objectives elicited by two different teams (A and B) are shown as an objectives tree (Figure 4.3) and as a list of objectives (Table 4.9) as they presented them in their final reports. While both lists may have some errors or problems that are worth noting, there are a number of interesting points to be made about these two sets of objectives. First, the two sets of

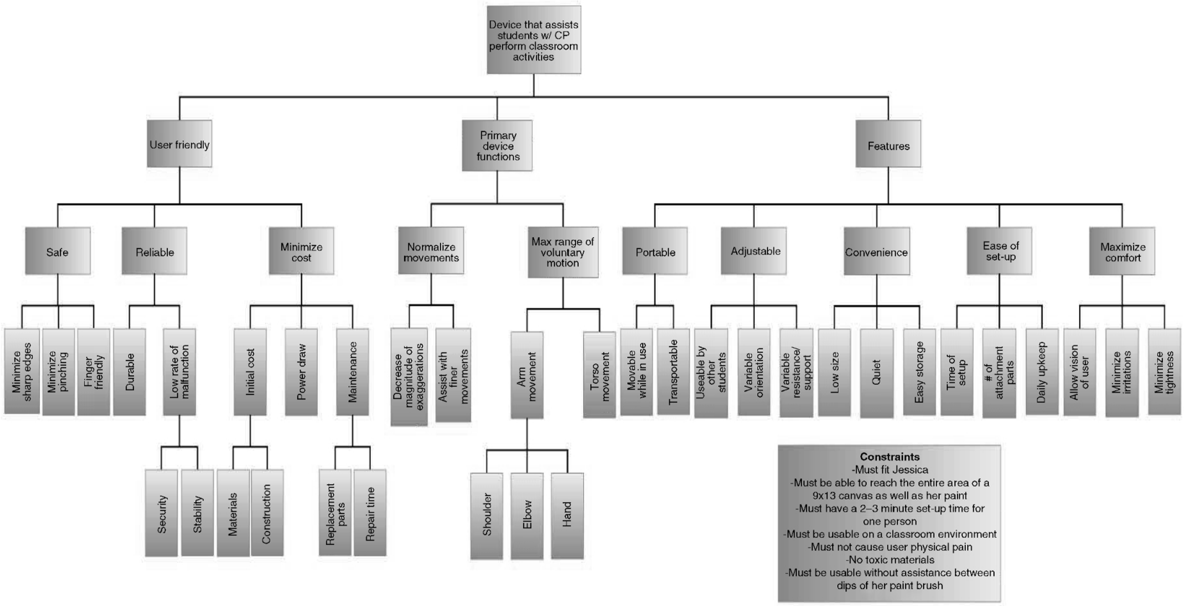



**Figure 4.3** Team A's objectives tree for the Danbury arm support project. Are there any entries in this tree that don't belong? Adapted from Attarian et al. (2007).

**TABLE 4.9 Team B's lists of objectives for the Danbury arm support project. How does this set of objectives compare with those shown in Figure 4.3? Are all of its entries appropriate?**

Design objectives

- Design should minimize involuntary movement of upper arm
  - Should be safe
  - Should be comfortable
  - Should be durable
  - Should not impair/restrict voluntary motion
- Design should be applicable to multiple individuals and wheelchairs
  - Size should be adjustable
  - Mounting mechanism should be adaptable
- Design should minimize the cost of production
- Restraint mechanism should be easy to install and maintain

objectives are different. This is not surprising, given that they reflect the work of two different teams; it highlights the fact that the objectives that designers discern are subject to analysis, interpretation, and revision. That is why it is very important that we, as designers, carefully review our findings with our client(s) before proceeding too far in the design process.

It is also worth noting that Teams A's objectives tree (Figure 4.3) has been given in great detail, so that we can traverse that objectives tree to readily answer, with some specificity, questions like *How?* and *Why?*. Such a dense objectives tree, however, raises some other questions: Do we need to develop metrics for each and every sub- and sub-subobjective on an objectives tree? How many of these subsidiary objectives (and metrics) should we consider when we later select a design from among a set of design alternatives?

A second point to note is that team A incorporated much more detail (Figure 4.3), perhaps already reflecting some additional research into potential designs, than did team B, whose objectives are much more general (Table 4.10), perhaps reflecting only what the client indicated in personal interviews.

Each team developed and applied metrics for its own set of objectives. Table 4.10 shows the metrics developed for 12 of the 23 objectives identified by team A; they illustrate some interesting issues. First, only *outcomes* are given: There are no scales or units. Rather, in their 62-page final report, the team wrote: "The specific units and scales for each metric are not presented due to size constraints." We might wonder what a client (or others) thinks about the apparent unwillingness to make room to credibly document the basis for design selection. In fact, team A developed some neat results (to be presented later), but their impact could be lessened because of such inattention to detail. Second, these metrics were actually developed not for the three objectives at the second level of team A's tree (Figure 4.3), or for the 10 subobjectives at the tree's third level, but for each of the 23 sub-sub-subobjectives displayed in the tree's fourth level. Thus, the higher-level objectives were not directly assessed, perhaps because they were sufficiently abstract that meaningful measurements could not be made. Third and last, some of the formal metrics appear to be very qualitative. It may often be that only qualitative assessments are possible, but a client will find it easier to accept such judgments when complete details are given for those objectives whose metrics can be measured.

Table 4.11 shows the metrics and their corresponding scales and units developed by team B. The metrics leave no doubt about what's to be tested, and how to measure the achievement of their objectives. This approach led to the rapid adoption of a very small set

**TABLE 4.10 This table shows metrics for 12 of the 23 objectives in the fourth level of Team A's objectives tree of Figure 4.3**

| Objectives | Metrics |
|---|---|
| 1. Minimize number of sharp edges | Number of sharp edges |
| 2. Minimize pinching | Number of pinching possibilities |
| 3. Finger friendly | Number of places on device to get finger caught |
| 4. Durable | Disconfiguration, misalignments of device after regular use |
| 5. Remain secure on user | Conditions under which device remains securely attached to user |
| 6. Maintain stable position | Conditions where position, orientation of device maintain mountain setting |
| 7. Minimize cost | Estimate dollar amount |
| 8. Normalize arm movement | User ability to draw straight line compared to do so without the device |
| 9. Maximize range of voluntary motion | Degree of freedom in motion of wrist, elbow, arm, and torso |
| 10. Movable while in use | Required assembly condition to move device |
| 11. Transportable | Necessary level of disassembly for movement |
| 12. Useable by multiple | Range of permissible arm sizes |

**TABLE 4.11 Team B's metrics for their objectives in Table 4.9**

| Objectives | Metrics |
|---|---|
| Safety | Measured by number of possible ways in which device can cause bodily harm.<br>Scale: Total Points = 10 − # ways to cause harm |
| Stabilization | Ability to resist sudden acceleration.<br>Scale: 1–10 by Subjective Evaluation |
| Comfortable | Perceived comfort of device.<br>Scale: Total Points = 10 − # sources of discomfort |
| Non-Restrictive | Measured by the area of allowed motion.<br>Scale: Total Points = 10 − (Area/2 sq. ft.) |
| Ease of Installation | Measured by the number of minutes required for installation.<br>Scale: Total Points = 10 − 2(Minutes Required) |
| Durable | Measured by flimsiness, points of failure, and ability to resist torques.<br>Scale: Total Points = 10 − # of points of failure |
| Adjustability | Measured by the device's ability to fit a range of wheelchairs and individuals.<br>Scale: 1–10 by Subjective Evaluation |
| Low Cost | Determined of the production cost of one unit.<br>Scale: Total Points: 10 − (cost/$200) |

**TABLE 4.12 Team B's pairwise comparison chart (PCC) for the Danbury arm support project**

| | Safety | Stability | Comfort | Durability | Non-Restrictive | Adjustable | Low Cost | Easy to Install | Total |
|---|---|---|---|---|---|---|---|---|---|
| Safety | – | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 7 |
| Stability | 0 | – | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
| Comfort | 0 | 0 | – | 1 | 1 | 1 | 1 | 1 | 5 |
| Durability | 0 | 0 | 0 | – | 0 | 1 | 1 | 0 | 2 |
| Non-Restrictive | 0 | 0 | 0 | 1 | – | 1 | 1 | 1 | 4 |
| Adjustable | 0 | 0 | 0 | 0 | 0 | – | 1 | 0 | 1 |
| Low Cost | 0 | 0 | 0 | 0 | 0 | 0 | – | 0 | 0 |
| Easy to Install | 0 | 0 | 0 | 1 | 0 | 1 | 1 | – | 3 |

of design alternatives, within which the selection of components remained somewhat larger. One minor quibble is that it would have been more helpful if the metrics were listed in the same order as their corresponding objectives in Table 4.9.

As a final note on the lists of objectives generated by teams A and B, we look at how they ranked or prioritized their individual lists. Team A did not include a pairwise comparison chart in their final report. Rather, they stated that, "The primary objectives were ranked on (1) the piecewise [their wording] comparison charts of the individual team members, in which all possible pairings of objectives were ranked, thus creating a complete ranking order, and (2) the recommendations of the team's liaisons." Team A's objectives rankings are thus reflected in the order in which they appear in the list of metrics in Table 4.10. Team B reported a formal PCC, which is shown in Table 4.12, and put that ordering to use when they evaluated their competing designs.

## 4.6 NOTES

*Section 4.1*: More examples of objectives trees can be found in Cross (1994), Dieter and Schmidt (2012), and Suh (1990).

*Section 4.2*: Measurements and scales are very important in all aspects of engineering, and not just design. Our discussion takes on a positivist approach (Jones 1992, Otto 1995).

*Section 4.3*: Some aspects of measurements have recently become controversial in the design community, to a degree beyond our current scope. Some of the critiques derive from an attempt to make design choices and methods emulate long-established approaches of economics and social choice theory (Arrow 1951, Hazelrigg 1996, 2001, Saari 1995, 2001a, 2001b). The PCCs outlined in the text are exactly the same as the best tool offered by the social choice theorists, the Borda count (Dym, Wood, and Scott 2003).

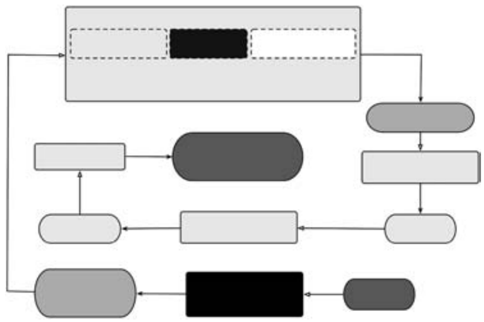
*Section 4.4*: Our discussion of metrics is strongly influenced by the German approach to design (Pahl and Beitz 1996).

*Section 4.5*: The results for the Danbury arm support design project are taken from final reports (Attarian et al. 2007, Best et al. 2007) submitted during the Spring 2007 offering of Harvey Mudd College's first-year design course, E4: Introduction to Engineering Design. The course is described in greater detail in Dym (1994b).

CHAPTER 5

# PROBLEM DEFINITION: IDENTIFYING CONSTRAINTS

*What are the limits for this design problem?*



WE CONTINUE our discussion of *problem definition* by focusing on identifying the *constraints* that must be satisfied, that is, by identifying limits that cannot be exceeded and boundaries that may not be crossed.

## 5.1 IDENTIFYING AND SETTING THE CLIENT'S LIMITS

Recall that in Chapter 3 we talked about questioning our client in order to better understand the problem. One of the questions we suggested asking was

- Are there things or circumstances you want us to avoid?

  This question might also have been phrased in terms of boundaries the client did not want crossed or limits that could not be exceeded, or numbers that were to be treated as "hard caps." Whatever the wording, we are talking about constraints:

- **constraint** *n*: a limit or restriction on the design's behaviors or attributes
  *Constraints* are typically framed as a binary yes-or-no choice: a ladder material is a conductor or it is not, or the step deflection is less than 0.05 in. or it is not. Any (and all) designs that violate these limits are unacceptable. Constraints are important to the design process because they limit the size of a design space by forcing the

**TABLE 5.1 The list of safe ladder constraints extracted from the list of attributes of Table 3.1**

| Characteristic | O | C | F | M |
|---|---|---|---|---|
| Step deflections should be less than 0.05 in. | | ✓ | | |
| Must meet OSHA requirements | | ✓ | | |
| Must not conduct electricity | | ✓ | ✓ | |

exclusion of unacceptable alternatives. For example, a ladder design that fails to meet OSHA standards must be rejected (Table 5.1).

There are limits to everything, of course, but as a practical matter we often use constraints as a kind of "checklist" to help us keep our list of possible designs to a reasonable length. Such constraints are typically expressed in terms of specific numerical values, but not always, as we can see from the safe ladder constraint list in Table 5.1. By way of contrast, objectives are much more likely to be expressed as verbal statements, for example, a ladder should be cheap.

Objectives and constraints are closely related, but they are not interchangeable. Constraints limit the size of the design space (i.e., the number of potential designs we might consider), while objectives permit us to explore what remains in that design space. Constraints enable us to reject unacceptable alternatives, while objectives enable us to select among design alternatives that are at least acceptable, or, in other words, designs that *satisfice*. Designs that satisfice may not be optimal or the best, but at least they satisfy all constraints. For example, we could minimally satisfy OSHA standards or we could significantly surpass those standards by making a "super safe" ladder in order to obtain a marketing advantage. Or, on the price side, a goal that a ladder should be "inexpensive" might also have a constraint that the ladder's cost cannot exceed $25. If we have *both* a low-cost objective and a $25 constraint, we may exclude some initial designs based on the constraint, while choosing among the remaining designs based on cost and other, non-economic objectives.

In Table 5.2 we list the constraints for the new juice container. Note that one constraint, chemically inert, is related to two of our objectives, safe and long shelf life.

## 5.2 DISPLAYING AND USING CONSTRAINTS

Constraints can be displayed in several ways. We can simply construct a list, as shown in Table 5.2. We can also add them to objectives trees, as we do in Figure 5.1 . When we add them to the objectives tree, we need to clearly distinguish them from the objectives

**TABLE 5.2 A list of constraints for the juice container design**

| |
|---|
| *Chemically Inert* |
| *No Sharp Edges* |

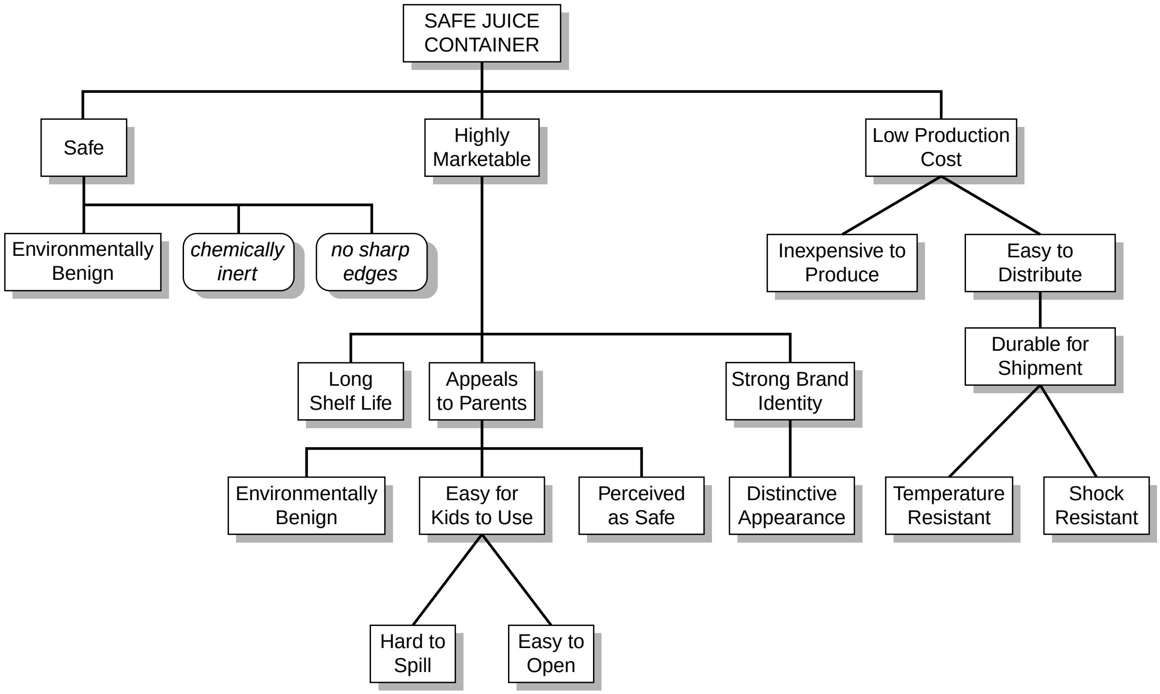



**Figure 5.1** A combined objectives (rectangles) and constraints (ovals) tree for the design of a new juice container. Compare this with Figure 4.2 that shows the "standard" objectives tree for the juice container.

(e.g., we present constraints in boxes differently shaped than those used for objectives, as in Figure 5.1), and also select an appropriate place for them. In the case of the juice containers, chemically inert may be placed under both safe and long shelf life. Similarly, if we wanted to present an outline form of an objectives–constraint hierarchy, we might enter the constraints in italics or a different font. In either case, it is most important to recognize that constraints are related to but are different from objectives: they mean very different things and are used in different ways. These combined objectives–constraints trees can be a very effective way to communicate how these two very different concepts, objectives and constraints, interact especially when talking to clients or to nontechnical audiences.

In addition to expressing the client's limits on the design, we will find that constraints are also useful later in the design process, both to help us prune or narrow our space of designs, and to help us do our screening and evaluation of designs.

## 5.3 CONSTRAINTS FOR THE DANBURY ARM SUPPORT

The constraints developed by team A for the arm support designed for Danbury School's Jessica are listed in Table 5.3, while those developed by team B are listed in Table 5.4. Particularly when placed in close proximity, these two constraints lists differ rather markedly. One difference is that team A's list has much more granularity in that it

**TABLE 5.3 Team A's list of constraints for the Danbury arm support project**

Design Constraints

- Must fit Jane
- Must allow Jane to reach the entire area of a 9 in. by 13 in. canvas as well as her paint, placed directly to the left side of the canvas
- Must have a setup time of 8 min or less
- Must be usable in a classroom environment
- Must not cause Jane physical pain
- Must not contain any toxic materials
- Must be usable without assistance between dips of the paintbrush during painting

**TABLE 5.4 Team B's list of constraints for the Danbury arm support project**

Design Constraints

- Design must reduce and counteract involuntary movement of the upper arm
- Design must not require more than 2–3 min for setup by an adult

provides much more detail about the meaning of many of the constraints and the environment is which the entire design is set. Some are easily envisioned as binary choices (e.g., team A's "Must not contain any toxic material"). Others sound fuzzier and perhaps more like objectives (e.g., team B's "Design must not require more than two to three minutes for setup by an adult"). Still more interesting is the fact that the teams seem to have identified different limiting values for the same constraint: team A says that setup must be "8 min or less," while team B says "not more than 2–3 min." Perhaps this difference arose because of the different conversations each team had with the Danbury School's teaching staff, thus reflecting some variation in what different teachers thought was a reasonable setup time. Perhaps it reflected an unease felt by the staff when they were asked for—or encouraged, or even pushed—to provide hard numbers. And perhaps they really meant short setup times as an objective, rather than a constraint.

It is also interesting that team A had identified an objective (Table 4.10) of "Minimize number of sharp edges," but that doesn't appear in any of their constraints, and we have to wonder why? Is it because this is a customized, "one of" design? Certainly, if we were designing a support arm to be marketed commercially, we would have to consider enforcing a constraint of having *no* sharp edges, rather than simply minimizing their number. In general, it is likely that the design of a widely used product will have to be done in a more constrained design space than client- or user-specific devices.

## 5.4 NOTES

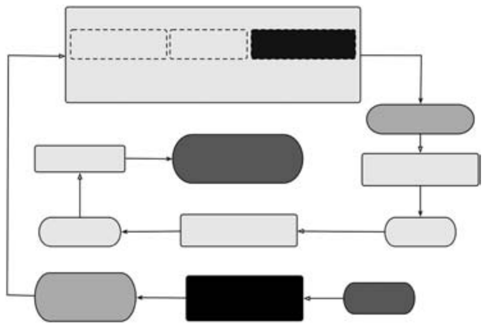
*Section 5.1*: Constraints are discussed in Pahl and Beitz (1996). The very important notion of *satisficing* is due to Simon (1981).

*Section 5.2*: The results for the Danbury arm support design project are taken from final reports by Attarian et al. (2007) and Best et al. (2007) submitted during the Spring 2007 offering of Harvey Mudd College's first-year design course, E5: Introduction to Engineering Design.

CHAPTER 6

# *PROBLEM DEFINITION: ESTABLISHING FUNCTIONS*

*How do I express a design's functions in engineering terms?*



WE NOW finish defining the client's design problem and move into engineering practice by (1) establishing the *functions* that the design must perform, and (2) writing *specifications* that express those functions in quantitative, engineering terms that enable us to ensure that those functions are performed. Since specifications detail or specify how the performance of those functions can be assessed, they have much in common with the metrics we use to assess the achievement of objectives, and very similar approaches can be taken to writing requirements for a design's features and behaviors.

## 6.1 ESTABLISHING FUNCTIONS

Asked what a bookcase does, a child might answer, "It doesn't *do* anything, it just sits there." An engineer, however, would say that the bookcase does at least two things: It resists the force of gravity exactly to support the weight of the books, and it enables the organization of those books with dividers or by its shelf lengths. Thus, this bookcase doesn't "just sit there." Understanding functionality is essential to successful design. There are consequences, often tragic, for failing to understand and design for *all* of a design's functions: recall the Hyatt Regency failure we discussed in Chapter 1. We will now explore how we talk about what designs do and then describe ways to establish functions.

### 6.1.1 Functions: Input Is Transformed into Output

We begin with our dictionary definition from Chapter 1:

- **function** *n*: those things a designed device or system is supposed to do.

For our work as designers, it is helpful to take a systems view and relate doing something to *transforming* an *input* into an *output*. Of course, this is also reminiscent of elementary calculus in which we write $y = f(x)$ to denote how the function $f(x)$ *transforms* the *input* of the independent variable $x$ into an *output* of a dependent variable $y$. For most of our purposes, *engineering functions* involve the transformation or transfer or flow of *energy*, *materials*, or *information*. We frequently view such transformations through the prisms of the conservation and balance principles that we detail when we discuss physical modeling in Chapter 9.

We see *energy* in mechanical, thermal, fluid, or electrical forms, and we see energy transformed as it is stored, transmitted, converted, or dissipated. We also view energy transformation or transfer to include forces transmitted or used to support (conservation of momentum or balance of forces), current flows (conservation of charge), and so on. We must account for all of the energy going into and coming out of a device or a system. This does not mean that the device or system is an ideal one in which energy is conserved. Rather, it means that energy can't simply disappear, even when it is dissipated.

Similarly, materials flow occurs in a variety of ways: moving or flowing through some conveyance (like a pipe), being transferred or located in a container, being separated into constituents, or added to, mixed in with, or located within one or more materials. Thus, cement, aggregate, and water are mixed to create concrete, which is then typically moved (while being mixed), poured, finished, and allowed to set and harden.

Finally, information flow includes the transfer of data in any of several forms: tables and charts on paper, data transmitted over the Internet or by wireless, and electrical or mechanical signals transmitted to sense or measure behavior or control response. The transformation of information occurs when, for example, a room temperature measured by a thermometer is transmitted to a wall thermostat that then instructs a heater or an air conditioner to change what it is doing. We might even think of the energy that is required to turn data into information and information into knowledge.

### 6.1.2 Expressing Functions

Given that functions are the things that a designed device must do, the statement of a function typically couples an *action verb* to a noun or object: *lift* a book, *support* a shelf, *transmit* a current, *measure* a temperature, or *switch* on a light.

The object in a verb–noun formulation function may start off with a specific reference to a particular design idea, but it is usually best to look for more generality. For example, while one bookcase function is "support books," bookcase shelves often support trophies, art, or even piles of homework. Thus, a more general, more useful statement of the function to be performed is to "resist forces due to gravity," which can in turn be associated with any objects weighing less than some predetermined weight (i.e., force): *support* a given number of kilograms (or pounds). When describing functions, then, we should use a verb–noun combination that best describes the most general case.

Similarly, we also should avoid tying a function to a particular solution. If we were designing a cigarette lighter, for example, we should avoid "applying flame to tobacco" because it eliminates car lighters that use electrical resistance in a wire. That specific formulation also eliminates using the lighter to ignite paper, wood, or charcoal briquettes.

We can also categorize functions as being either *basic* or *secondary* functions. A *basic function* is the specific, overall function that must be performed, and *secondary functions* are (1) other functions needed to perform the basic function or (2) those that result from doing the basic function. Secondary functions maybe categorized further as either required or unwanted functions. *Required secondary functions* are those needed for the basic function. For example, the basic function of an overhead projector is to project images. This requires several secondary functions, including converting energy, generating light and focusing images. The projector also produces *unwanted secondary functions* such as generating heat and generating noise. Of course, such undesirable by-products may also spawn new required functions, for example, quieting noise or dissipating generated heat. This last example also suggests that we should try to anticipate all secondary functions, lest they turn into undesirable *unanticipated side effects* that may significantly affect how a new design is perceived and accepted.

## 6.2 FUNCTIONAL ANALYSIS: TOOLS FOR ESTABLISHING FUNCTIONS

We now provide several tools to do *functional analysis*, that is, to establish the functions that our design must perform. Our starting point will be the notion of the design as transformer of inputs to outputs: We define the boundary between our device and its surroundings, and then examine the inputs to and outputs from the device that cross that boundary. We (1) track the flow of energy, materials, or information through the device's boundary, and (2) detail how those inputs are used, converted, or otherwise processed to produce the desired functions. The device itself, contained within the specified boundaries, may be a "black box" if we've no idea what's inside, or a "transparent" box if we do. We will also describe three other tools used to establish functions: enumeration, dissection or reverse engineering, and function–means trees.

### 6.2.1 Black Boxes and Glass Boxes

One tool for understanding the connections between inputs and outputs is the *black box*: a graphic of the system or object being designed, with inputs shown entering the box on its left-hand side and outputs leaving on the right. *All* of the known inputs and outputs should be specified, even undesirable byproducts that result from unwanted secondary functions. For example, consider the black box for a power drill shown in Figure 6.1. We can think of the power drill (system) as a box that transforms the controlled power input into a rotating chuck, in which we can insert a drill bit to drill a hole or a screwdriver blade to drive a screw. At the top level, we can think of a power drill as having just three inputs: a source of electrical power (electrical energy), a supporting force (mechanical work) that holds or

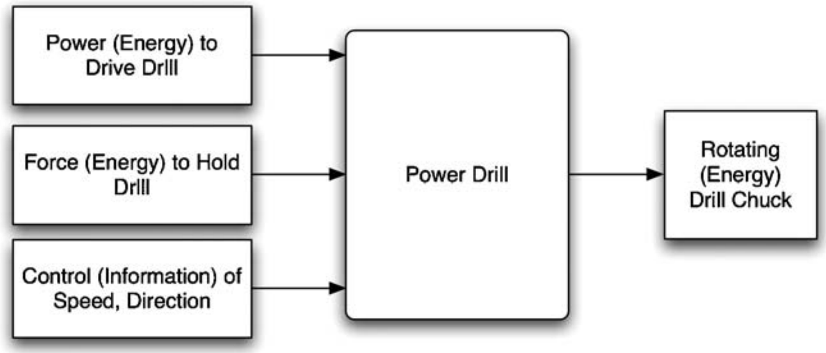



**Figure 6.1** A "black box" for a power drill with inputs and the output is encapsulated in one *basic function*: provides power to a screw. We take the cover of this black box in Figure 6.2.

grasps the drill, and the control of speed and direction (information) of the drill chuck's rotation. The drill has two outputs: a rotating chuck and a force holding the drill. (We will ignore heat generation and losses in this example.)

A top-level black box such as that shown in Figure 6.1 almost certainly prompts more questions than answers: How does this actually happen? What functions are performed in a power drill? Can we identify all of the (many) subfunctions performed inside of the power drill's black box? We begin to answer these questions by taking the cover off this black box and turning it into a *glass box* or a *transparent box* (Figure 6.2). (We will see another set of

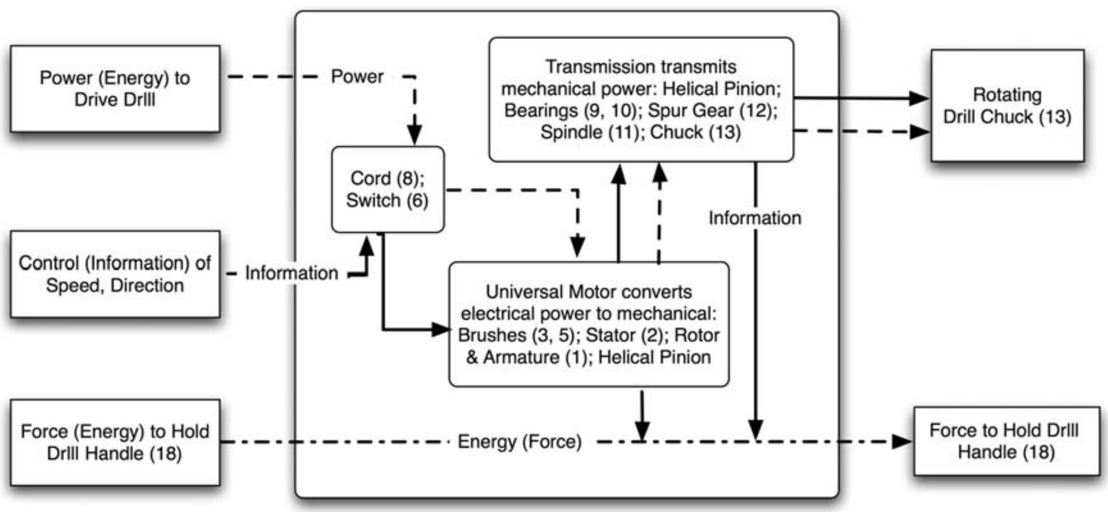



**Figure 6.2** The cover on the black box of Figure 6.1 has been removed, "drilling down one level" to make it transparent (i.e., a glass box), thus exposing a large number of secondary functions that are needed to perform the top-level functions. Compare this with the exploded view of the drill in Figure 6.3.

answers to these questions when we dissect this drill later in the chapter.) The glass box exposes several new boxes within the original (black) box, each of which represents a subfunction that must be performed to support the drill's overall functionality. These new boxes include a power cord (8) transmitting electrical power into the drill, where a switch (6) both directs that power and transmits information about its level to a universal motor. The universal motor converts the electrical power into mechanical power and transmits that power and information about its level to the transmission. The transmission increases the torque output by, in this particular drill, reducing the speed of power transmission from the 30,000 rpm of the rotor (1) in the universal motor to a (peak) output of 2500 rpm for the rotating chuck (13). Note that even in this brief description, it is often difficult to describe everything that happens at the same level of detail. If we were assembling this drill from known components or parts, we might have enough information to stop at this level. The recurring theme is that black boxes are made transparent when we ask *how* inputs are transformed into outputs, to the level that we need to fully achieve the top-level functionality of our design.

It is important that we *reflect the underlying physics* and express the inputs in those terms when we're applying the black-to-glass box approach. In the drill example, we didn't just hold and point the drill, we provided a force to hold it and information to both direct it and control its speed. Similarly, if we were doing this exercise for an old-fashioned radio, we would recognize that tuning a radio means selecting a particular frequency in the radio frequency (RF) spectrum, not just turning a dial.

The black-to-glass box idea can also be used effectively for systems or devices that do not have a *physical* box or housing. The only requirements for using black and glass boxes are that the device's boundary is specified and *all* of the inputs and outputs are identified.

We must be careful when we set the *boundary* of a system or subsystem whose functions we are identifying with the black-to-glass box tool because there is a trade-off. If we set boundaries too broadly, we may incorporate functions that are beyond our control, for example, generating household electric current for the drill or radio. If we draw boundaries too narrowly, we may limit the scope of the design. For example, the radio output could be an electrical signal that is fed to speakers, or it could be the acoustic signal coming from speakers. So, is the boundary drawn to include the speakers, or not? Such decisions should have been resolved during problem framing, when the problem's scope was set.

### 6.2.2 Dissection or Reverse Engineering

Our second functional analysis tool reflects the curiosity that most engineers feel. When confronted with a button, knob, or dial, they ask, "What does this do?" And they follow up with "How does it do that?" or "Why would you want to do that?" Questions like these, as well as following thoughts on how we might do it better or differently, form the core of *dissection* or *reverse engineering*. We *reverse engineer* a device or system that does some or all of what we want our design to do when we take it apart: We *dissect*, deconstruct, or disassemble it. We do this to find out in detail just how it works so we can then apply what we learn to our own design problem. We may not be able to use that design for any number of reasons: It may not do all the things we want, or may not do them very well; it may be too expensive; it may be protected by a patent; or it may be our competitor's design. But even if all of these reasons apply, we often can gain insight into our own design problem by looking at how others have thought about the same or similar problems. (Remember: Design problems are *open-ended* because they typically have several acceptable solutions!)

Reverse engineering is, in principle, a simple process. We look at the parts (e.g., gears, levers, circuit elements) that are used in a design or a device, and ask what functions those parts perform. We then look for alternative ways to do the same thing(s). For example, a button on an overhead transparency projector, when pushed, turns on the projector. To perform the function of turning the projector on or off, we might consider toggle switches or bars along the front of the projector.

We showed in Figures 6.1 and 6.2, respectively, black and glass boxes for a considerably more complicated device, a power drill. In Figure 6.3 we show an exploded view of a DeWalt™ corded power drill (Model D21008K). The individual pieces in that exploded view are the physical counterparts of the major subsystems we detailed in the glass box. We see an actual power cord (8) and a switch (6) connected to pieces of a motor (3, 5, 2, 1) and of a transmission (9, 10, 12, 11, 13). We ask what these pieces do, and find that the cord transmits electric energy, the switch controls the input of that energy, the motor turns the electrical energy into mechanical energy, which the transmission uses to control the drill's torque output. This "visual" inspection, actually the dissection or reverse engineering of the drill, produces the same functional analysis that we found with the black-to-glass box analysis.

We have already noted that there are often good reasons why we can't use a particular device or design we're dissecting. First, that device or design was developed to meet the goals of a particular client and a target set of users, and they may have had different concerns than we have. We have to remember to stay focused on our client's project. Second, adapting a new subfunction or new means for a function from the dissected device might limit how we think about our project. For example, if we become captive to a really nifty switch for turning on the power to a stand-alone overhead projector, we might not think about incorporating a projector's controls with a room-based approach, where the controls are integrated and mounted in a wall panel or on a lectern or a keyboard. This is a reinforcement of what we said earlier: We should try to define functions as broadly as we can, and with as much attention to the physics as possible. If we restrict ourselves to the most immediate expression of functions found in someone else's design, we limit our creative possibilities—and we may well run into serious intellectual property and ethical issues.

Finally, while we treat the terms reverse engineering and dissection as equivalent in the context of engineering design, in other fields dissection may be seen as purely descriptive: We dissect something just to reveal its underlying physical structure. In reverse engineering, we try to be analytical as we look at the same physical structure in order to identify the means to make functions happen, that is, we are trying to analyze both the functions of a device and how those functions are (and can be) implemented.

### 6.2.3 Enumeration

Another basic method of determining functions for a designed object is to simply *enumerate* or list all of the functions that we can readily identify. This is an excellent

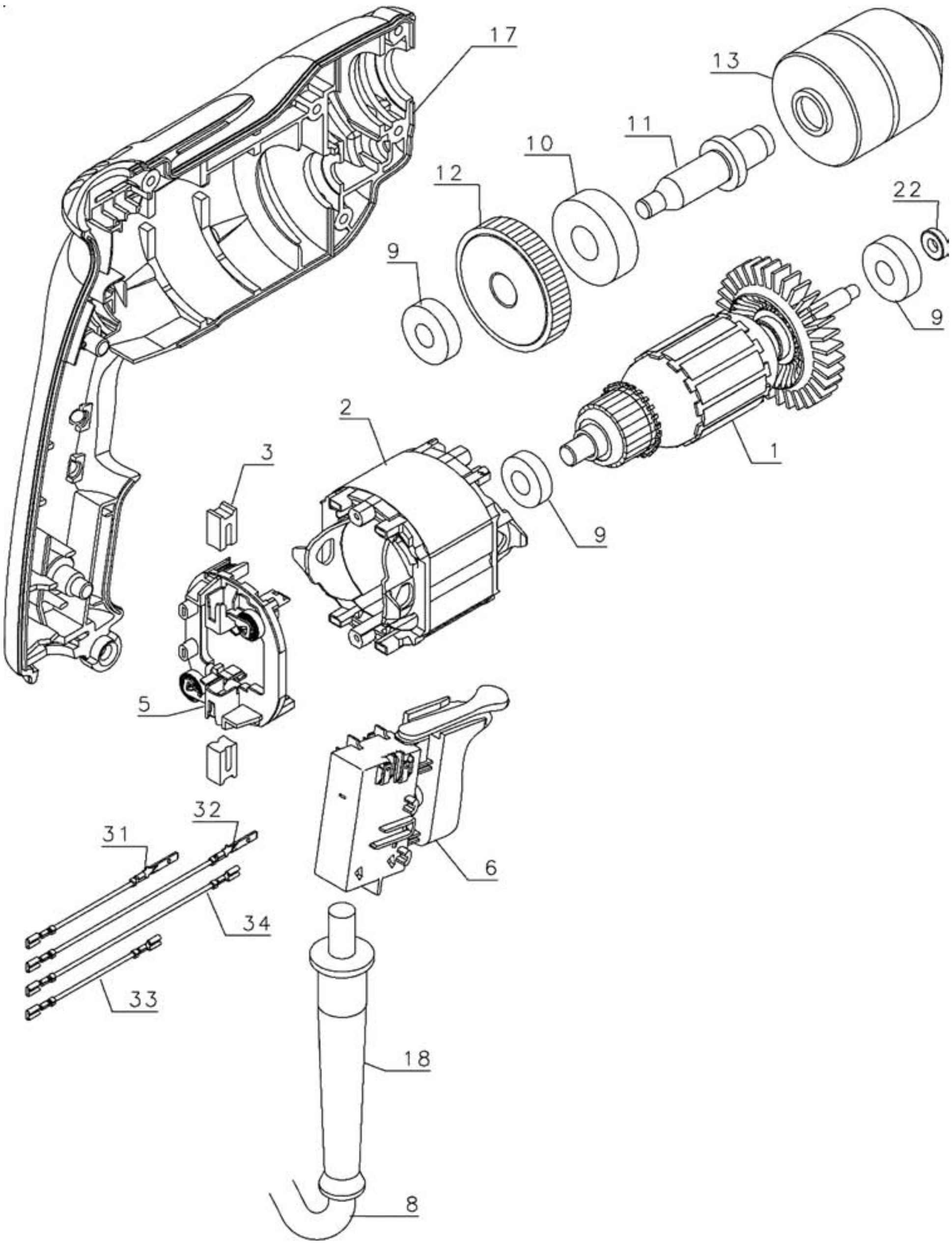



**Figure 6.3** An exploded view of the major subsystems in a DeWalt™ corded power drill, Model D21008K, showing several major subsystems: *switch* (6); *universal motor* consisting of brushes (3, 5), stator (2), rotor and armature (1), helical pinion; *transmission* consisting of helical pinion, bearings (9, 10), spur gear (12), spindle (11) and chuck (13); and the *clam shell* cover in the upper left. Courtesy of *Black & Decker Corporation*.

way to begin functional analysis because it leads us immediately to the basic function(s) of the device. It may be problematic for determining secondary functions, however, because a lot of specific engineering background might be required. Consider the drill in Figure 6.3: It would hard to be much more specific than our black box for the drill (Figure 6.1) without knowing about gears, motors, switches, and so on. Thus, we might get "stumped" very early in this process. However, enumeration is more than just making a list. Successful enumeration requires thoughtful questioning—much in the spirit of design as questioning—so we now provide some useful "tricks" to extend our enumerative thinking.

One such trick is to imagine that object *X* ceases to exist by asking, "What happens if there is no *X*?" If a bridge disappeared entirely, for example, any cars on the bridge would fall into the river or ravine over which the bridge crosses. This suggests that one function of a bridge is to support loads placed on the bridge. If the abutments ceased to exist, the deck and superstructure of the bridge would also fall, which suggests that another function of the bridge is to support its own weight. (This may seem silly until we recall that there have been more than a few disasters in which bridges collapsed because they failed to support even their own weight as they were being built! One of the most famous of such infelicitous bridge designs is the Quebec Bridge over the St. Lawrence River: It collapsed in 1907 with a loss of 75 lives, and again in 1916 when its closing span fell down.)

Similarly, if the connections of a bridge to various roads disappeared, traffic would not be able to get on the bridge, and vehicles on the bridge could not get off. This suggests that another function that a bridge serves is to connect the river/ravine crossing to a road network. If the bridge's road dividers were deleted, vehicles headed in one direction could collide with vehicles headed in the other. Thus, another bridge function is to separate traffic, a function that can be accomplished in several ways: New York's George Washington Bridge puts different traffic directions on different levels. Other bridges use median strips.

Another way to determine an object's functions is to ask how it might be used and/or maintained over its life. For example, to maintain our bridge we'd want to provide access to both inspectors and painters, which could be done with several means (e.g., ladders, catwalks, elevators). Similarly, we might think about the lifecycle of our juice container design. We can easily list the functions performed by a juice container because we use them all the time:

- contain liquid;
- get liquid into the container (fill and seal the container);
- get liquid out of the container (empty the container);
- close the container after opening (if it is to be used more than once);
- resist forces induced by temperature extremes;
- resist forces induced by handling in transit; and
- identify the product.

We note that the functions of containing liquid and of filling and emptying the container are distinct, which is consistent with our experience: Liquid is poured in before the container is sealed by a permanent top, and emptying may be enabled by a pull tab. These functional distinctions (i.e., between filling, containing, and emptying) emerge when we considered the container's lifecycle.

At the heart of our approaches to function enumeration lie the dual needs to ask thoughtful questions and to properly use verb–noun pairs to express each and every function of the designed object.

### 6.2.4 Function–Means Trees

No matter how much we are warned against "marrying your first design" or cautioned against trying to solve design problems before we fully understand them, we often jump to early design ideas because of the functions we do see. Consider a handheld lighter. Here we use a flame to ignite leafy materials (i.e., tobacco), as opposed to an auto's lighter, which uses resistance heating. The handheld lighter likely requires that we shield the flame against wind, a secondary function we do not need for the auto's lighter. A function–means tree is a tool that helps us sort out secondary functions in cases where different means or implementations can lead to different subfunctions.

A *function–means tree* is a graphical representation of a design's basic and secondary functions. The tree's top level shows the basic function(s) to be met. Each succeeding level alternates between showing:

- the *means* (in trapezoids) by which the primary function(s) might be implemented, and
- the *secondary functions* (in rectangles) necessitated by those means.

We show a function–means tree for a handheld lighter in Figure 6.4. Note that the top-level function has been specified in the most general terms possible. At the next level, a flame and a hot wire are given as two different means. These two means imply different sets of secondary functions, as well as some common ones. Some of these secondary functions and their possible means are given in lower levels.

Once a function–means tree has been developed, we can list all of the functions that have been identified, noting which are common to all (or many) of the alternatives and which are particular to a specific means. Functions that are common to all of the means are likely to be inherent to the problem. Others are addressed only if the associated design concept is adopted after evaluation.

A function–means tree is also useful because it begins the process of relating *what* we must do to *how* we might do it. We will return to this issue in Chapter 7 when we introduce the *morphological chart* as a tool to help us generate design alternatives. The "morph" chart lists the functions of the designed device and the possible means for realizing each function in a matrix format. The effort we put into the function–means tree (now) will really pay off (then)!

As with all tools, function–means trees must be used carefully. First, a function–means tree *is not* a substitute either for framing the problem or for generating alternatives. It may be tempting to use the outcome of the function–means tree as a complete description of our available alternatives, but we will almost certainly restrict the design space much more than we have to. Second, a function–means tree should not be used alone, that is, without using some of the other tools described above. Sometimes people adopt a tool because it somehow "fits" with their preconceived ideas of a solution, a mistake akin to "marrying your first design" or reinforcing preconceived ideas. Because the function–means tree allows us to work with appealing means or

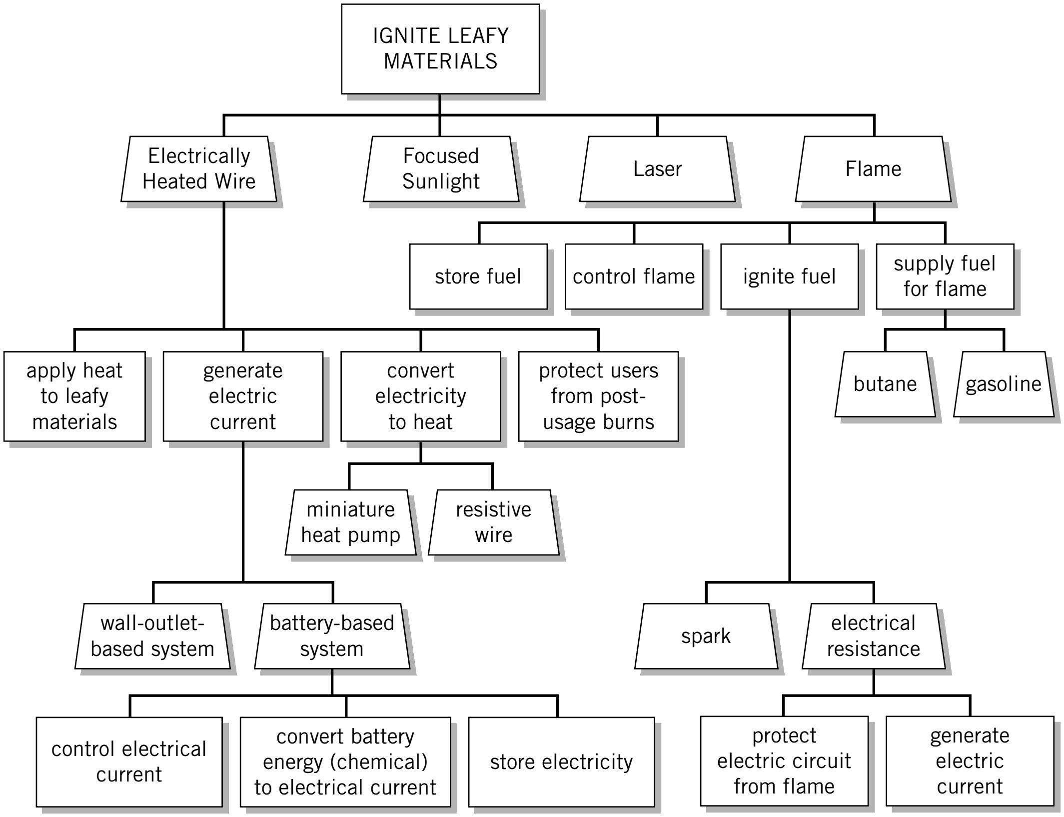



**Figure 6.4** A function–means tree for a lighter: functions in rectangles, means in trapezoids. Note that means produce different subfunctions.

implementations, we may overlook functions that might have turned up with a less "solution-oriented" technique.

### 6.2.5 Remarks on Functions and Objectives

It is often easy to mistake objectives for functions and functions for objectives. This happens because objectives sometimes express a functional need. For example, a bookshelf design might have as an objective that it hold the complete *Harry Potter* and *The Lord of the Rings* series of fables, whereas the specifications include both a function (i.e., support the weights of those collected sets) and a requirement that the shelf has the feature of being long enough to accommodate all 10 volumes in hardback editions (7 of *Harry Potter*, 3 of *The Lord of the Rings*).

We pointed out in our discussion of objectives trees in Section 4.1.2 that we are nearing the end of the objectives tree when the *why* question turns toward *how*, meaning that functions may begin emerging as ways to achieve objectives. We can reduce the confusion between objectives and functions if we keep in mind that objectives are expressed as *are* or *being* adjectives while functions are expressed as active *do* or *doing* verb–noun pairs. In fact, as we noted in our discussion of the objectives tree for the juice container in Section 4.1.3, another way that we might identify functions is by examining and rewriting detailed subobjectives because they suggest functions. So the subobjective Shock Resistant suggests the function Resist Shock. Recognizing the similarities and the differences of objectives and functions is quite important, as a lot of serious design practice will confirm.

## 6.3 DESIGN SPECIFICATIONS: SPECIFYING FUNCTIONS, FEATURES, AND BEHAVIOR

We noted in Section 6.1 that determining the functions of a designed object or system is essential to the design process. Functional requirements don't mean much if we don't consider *how well* a design must perform its functions. For example, if we want a device that produces musical sounds, we must specify how loudly, how clearly, and at what frequencies the sounds are produced. Earlier, in Chapter 1, we had noted that design specifications or requirements specify in engineering terms a design's functions, as well as its features and behaviors. Such requirements, often called "specs," provide a basis for determining a design because they are the targets against which we measure our success in performing or achieving them. Design specifications are presented in three forms that represent different ways of formalizing a design's functional performance and its features and behaviors for engineering analysis and design:

- *Prescriptive specifications* specify ***values** for attributes of design*. Thus, for our children's juice container we might say, "A juice container must be made of at least 50% recyclable plastic." For a ladder design we might say that "A ladder step shall be made from Grade A fir, have a thickness of at least 0.75 in., have a length that does not exceed 80 in., and be attached to the side rails through a full-width groove at each end."
- *Procedural specifications* specify ***procedures** for calculating attributes*. So we would say: "The juice container must be disposable as stipulated by EPA standards." In a ladder design we might specify: "The maximum bending stress $\sigma_{\max}$ in a ladder step shall be calculated from $\sigma_{\max} = Mc/I$ and shall not exceed the allowable stress $\sigma_{\text{allow}}$."
- *Performance specifications* specify ***performance levels** that a function must demonstrate to be successful*. Then we would say: "A juice container must contain 75 ml." And for the ladder project we'd say: "A ladder step shall support an 800 lb gorilla."

In addition, if a system or device has to work with other systems or devices, then we must specify how those systems interact. We call these particular requirements *interface performance specifications*.

### 6.3.1 Attaching Numbers to Design Specifications

It is normally the designer's job to express functions in engineering terms so that engineering principles can be appropriately applied to the design problem at hand. As designers, we have to cast functions into terms that enable us to measure how well a design realizes a specific function. That means that we have to establish both a range over which a measure is relevant to our design and the extent to which ranges of improvements in performance really matter.

When we talk about measuring the performance of a function, we are describing something that is conceptually the same as the *metrics* we introduced to measure the achievement of objectives. The thinking is very similar, so some of what we detailed in Chapter 4 about constructing metrics might be useful in this context, as what we say below is also relevant to thinking more deeply about objectives. But there are some key differences:

- We have reserved the term *metrics* to apply only to scaling *objectives*.
- We need a *metric for each* of the *objective*s that we will use in evaluating our design choices.
- We use *specifications* to scale *functions*.
- We need a specification for every function our design must perform, and designs must meet *each and every* specification. In fact, in some sense being fully functional can be considered a constraint on the design.
- Metrics are applied in the *past* tense, to assess whether objectives *have been* achieved.
- Specifications, such as constraints, are projected into the *future*, when we specify in advance the functional or behavioral performance that *must be* achieved in order for a design to be considered to be successful.

How do we determine the range over which a measure is relevant to a design and decide how much improvement is worthwhile? Our conceptual starting point is similar to what economists call a *utility plot* (see Figure 6.5(a)): It graphs the usefulness of an *incremental* or *marginal* gain in performance against the level of a particular design variable.

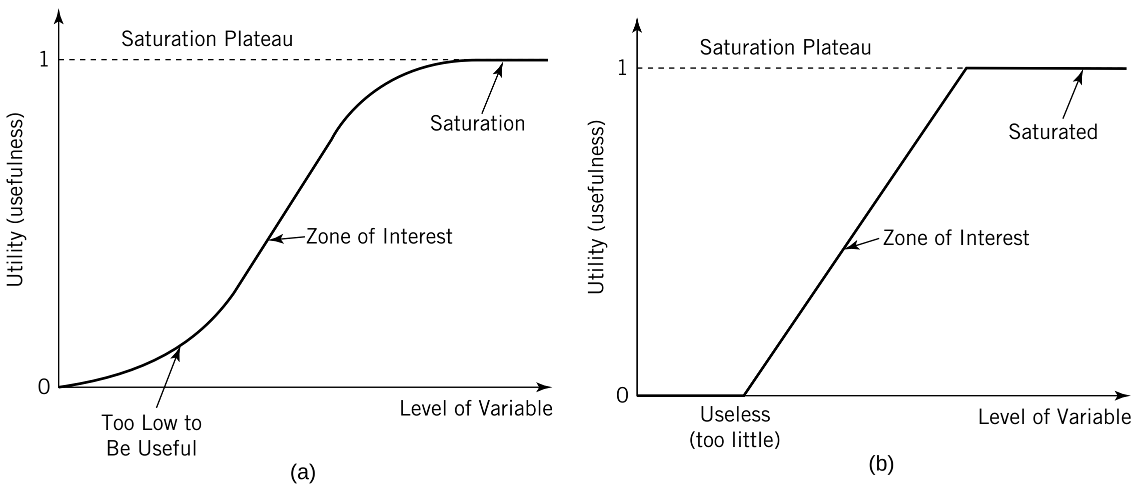



**Figure 6.5** Saturation curves showing that no additional benefit is achieved below some minimal realized level and above saturation. The actual shape of the hypothetical curve (a) is likely to be uncertain in most cases, while the linear approximation (b) is a design team's effort to make the meaning more useful by making the curve crisper.

We usually plot the utility or value of a design gain as the ordinate ($y$-axis) and normalize it to the range from 0 to 1. We plot the level of the attribute being assessed on the abscissa ($x$-axis). For example, consider using processor speed as a measure of a laptop computer performance. At processor speeds below 1 GHz, the computer is so slow that a marginal gain from, say, 500 to 750 MHz provides no real gain. Thus, for processor speeds below 1 GHz, the utility is 0. At the other end of the utility curve, say, above 5 GHz, the tasks for which this computer is designed cannot exploit additional gains in processor speed. For example, browsing the World Wide Web may be more constrained by typing or communication line speeds, so that an incremental gain from 5 to 5.1 GHz doesn't change the normalized utility of 1. Thus, the utility plot is *saturated* at high speeds.

What happens at performance levels between those that have no value and those at or above saturation, say in the range 1–5 GHz for the computer design? We expect that changes do matter in this range, and that we will see incremental or marginal gains with increases in processor speed. In Figure 6.5(a) we show an *S*- or *saturation* curve that displays *qualitatively* the gain achieved as a (qualitative) function of the processor speed: No numerical values are attached to either axis. Thus, while no specific value of gains can be determined from the curve, the *S*-curve demonstrates zero utility at low processor speeds, shows a measurable increase over a range of processor speeds, and then plateaus (or saturates) at 1 because increased processor speeds produce no gains in utility.

The sort of behavior seen in a utility plot is rather common. We don't always know the precise details of the *S*-curve: It may not look nearly as smooth as what we have sketched in Figure 6.5(a), so we approximate it by a set of straight lines, as we show in Figure 6.5(b). We still see regions where gains have no utility or are unavailable, as indicated by the horizontal lines at levels 0 and 1. In the range of interest, however, we model our utility level as a linear function of the design variable (e.g., processor speed). We are simply saying, qualitatively, that the linear curve defines a range within which we can expect to achieve gains by increasing the relevant design variable and, conversely, reduce gains by decreasing that design variable.

To take another example, suppose we are asked to design a Braille printer that is quiet enough to be used in office settings. None of the competing designs are quiet enough to be so used. How quiet does this design really have to be? To answer this question, we must determine the relevant units of noise measurement and the range of values of these units that are of interest. We should also find out how much noise is generated by current printer designs and whether listeners can distinguish different designs by their noise levels. If one printer produces the same noise level made by a pin dropped on a carpet, while another generates the noise level of a ticking watch, we would likely view both as quiet enough to be fully acceptable. Similarly, if one printer is as loud as a gas lawn mower, and another as an unmuffled truck, we gain no utility by distinguishing between these two designs as neither would be used in an office. (Note that this example shows a *reverse S*-curve in which we start at saturation because there is no gain to be made at such low levels of quietness, and then we degrade to a level of no utility for printers that are uniformly too loud.)

Since we measure sound intensity levels in decibels (dB), we know that *some* dB range is likely to be of interest, but what range? We answer this question by seeing how much noise is produced by other devices, and within different environments. We show sound intensities for various devices and environments in Table 6.1. With such environmental and exposure information in hand, we can identify a range of interest for a

**TABLE 6.1 Sound intensity levels that are produced by various devices and are measured in various environments. Sound intensity levels are measured in decibels (dB) and are a logarithmic expression of the square of acoustic power. Thus, a 3 dB shift corresponds to a doubling of the energy produced by the source, while the human ear cannot distinguish between levels that differ by only 1 dB (or less)**

| Level (dB) | Qualitative Description | Source/Environment |
|---|---|---|
| 10 | Very Faint | Hearing Threshold; Anechoic Chamber |
| 20 | Very Faint | Whisper; Empty Theater |
| 30 | Faint | Quiet Conversation |
| 50 | Faint | Normal Private Office |
| 50 | Moderate | Normal Office Background Noise |
| 60 | Moderate | Normal Private Conversation |
| 70 | Loud | Radio; Normal Street Noise |
| 80 | Loud | Electric Razor; Noisy Office |
| 90 | Very Loud | Band; Unmuffled Truck |
| 100 | Very Loud | Lawn Mower (Gas); Boiler Factory |

After Glover (1993).

performance specification for the Braille printer. New printer designs must generate less than 60 dB of noise in an office, and lower generated noise levels are considered gains, down to a level of 20 dB. All designs that generate less than 20 dB are equally good. All designs that produce more than 60 dB are unacceptable. Note that the printer noise levels we're talking about here are well below the limits that OSHA, the U.S. Occupational Health and Safety Administration, prescribes for occupational safety.

### 6.3.2 Setting Performance Levels

To set performance levels, we first identify design variables that reflect the functions that must be performed and the units in which those variables are measured. Then, assuming a standard or linearized *S*-curve, we establish the range of interest for each design variable by identifying the following: a *threshold* below which no meaningful gains can be made; a *saturation plateau* above which no useful gains can be achieved; and a range-of-interest zone that lies between the threshold and the plateau. It is within this zone that we map the design gains onto the design variables that are the subject of our performance specification. This works well if we exercise judgment in setting performance requirements based on sound engineering principles, an understanding of what can and cannot be reasonably measured, and an accurate reflection of both client's and users' interests.

Consider once again the juice container. Each of the functions specified in Section 6.2.3 has a range of values that must be specified. Some of those functions and some relevant questions associated with each function are as follows:

- *Contain liquid*: How much liquid must the container contain, and at what temperatures? Is there a range of fluid amounts that we can put into a container and still meet our specification?

- *Resist forces induced by temperature extremes*: What temperature ranges are relevant? How might we measure the forces created by thermal stresses on the container designs?
- *Resist forces induced by handling in transit*: What is the range of forces that a container might be subject to during routine handling? To what degree should these forces be resisted in order for the container to be acceptable?

Note that similar but distinct problems arise for the second and third functions on this list, as they both relate to forces.

We can now specify a set of performance specifications that the container designs should meet by addressing these and similar questions. For example, we might indicate that each container must hold $12 \pm 0.01$ oz. In this case our requirement has become a constraint because the corresponding utility plot is a simple binary switch: Either we meet this design specification or we don't. (Of course, we could study the container design problem as one in which the variable is a single-size serving, in which case we may find a reverse *S*-curve showing that a smaller container is better.) We might generate another performance specification by insisting that the containers can be filled by machines at a rate of 60–120 containers per minute (cpm). The threshold of 60 cpm might derive from the volume of sales needed to penetrate a market, while a saturation plateau of 120 cpm faster rate might exceed demand projections.

We might also specify that the designs should allow the filled containers to remain undamaged over temperatures from $-20$ to $+150\,^{\circ}$F. Temperatures lower than a threshold of $-20\,^{\circ}$F are unlikely to be encountered in normal shipping, while temperatures higher than a $+150\,^{\circ}$F plateau indicate a storage problem. It may be that some designs that appeal in other ways are limited here by either temperature extreme. We have to make a judgment about the importance of this function and its associated performance specification.

We also note that manufacturers or distributors often publish a product's performance specification *after* the product has been sent to market because users and consumers want to know whether the product is appropriate for *their* intended use. End users, however, are usually not parties to the design process, and so they depend on published performance specifications to know what they can expect from a product. In fact, designers often examine the performance specifications of similar or competing designs, in a manner much like reverse engineering, to gain insight into issues that may affect end users.

### 6.3.3 Interface Performance Specifications

As we noted earlier, we use *interface performance specifications* to detail how devices or systems must work together with other systems. These specifications are particularly important in cases where several teams of designers are working on different parts of a final product, and all of the parts are required to work together smoothly. For example, the design of a car radio must be compatible with the space, available power, and wiring harness of the car.

With such complexity, boundaries between subsystems must be clearly defined, and anything (i.e., energy, material, information) that crosses a boundary has to be specified in sufficient detail to allow all teams to proceed. These specifications may be a range of values of variables, physical or logic devices that support a boundary, or perhaps they are simply an agreement that a boundary cannot be breached (say between subsystems). In each case, designers of artifacts on both sides of a boundary must have a common understanding about where the boundary is and how it might be crossed, if at all. Reaching such understanding is hard because teams on both sides of a boundary are, in effect, placing constraints on their counterparts on the other side. A black-to-glass box analysis can be helpful in developing interface requirements because it allows all of the parties to identify the inputs and outputs that must be matched and so deal with any side effects or undesired outputs.

Interface performance requirements are increasingly important for large firms that, in a highly competitive international arena, are trying to minimize the total time needed to design, test, build, and bring to market new products. Most of the world's major automobile companies, for example, have reduced their design and development times for new cars to less than one-half of what they were a decade ago by having design teams work *concurrently*, or in parallel, on many subsystems or components. This means the teams must work together, and it puts a premium on understanding and working with interface performance issues.

### 6.3.4 House of Quality: Accounting for the Customers' Requirements

We now turn to a twofold task, ensuring a design's quality, and as part of that, ensuring that we've paid proper attention to what the customer would like to see in our design. What do we mean by *design for quality*? One answer is simple: *Quality* is "fitness for use," that is, quality is a measure of how well a product or service meets its specifications and requirements. By this definition, much of the problem framing activities discussed heretofore are aimed toward a "quality" design that meets or exceeds objectives, satisfies all constraints, and is fully functional—as well as or better than alternative designs. In that sense, all of our conceptual design work is directed to design for quality.

We also note that while detailing the attributes (i.e., the functions, features, and behaviors) so far, we have taken great care to stay aware of what the client wants, but we haven't paid much attention to what users might need or want in a product. One of the most important notions used by many designers is *quality function deployment* (QFD), which is expressed in a tool called the *House of Quality* (HoQ). A HoQ is a matrix that combines a lot of information about stakeholders, desirable characteristics of designed products, current designs, performance measures, and trade-offs. We show the general structure of a HoQ in Figure 6.6; it clearly shows a metaphorical house. The HoQ's *Who* refers to the stakeholders in the design process: client(s), users, and other affected parties. The *What* entries correspond to the design's desired attributes (i.e., objectives, constraints, functions) objectives. The *Now* entries refer to existing products or designs that are typically found during problem definition, and they are used for benchmarking proposed designs. The *How* elements refer to the metrics for objectives and the specifications for functions. The entries *How Much* or *Target* are goals or targets for the *What* entries. The remaining sections are devoted to the relationships, values, or trade-offs among the elements we've just described. These sections become much clearer if we look at a specific example.

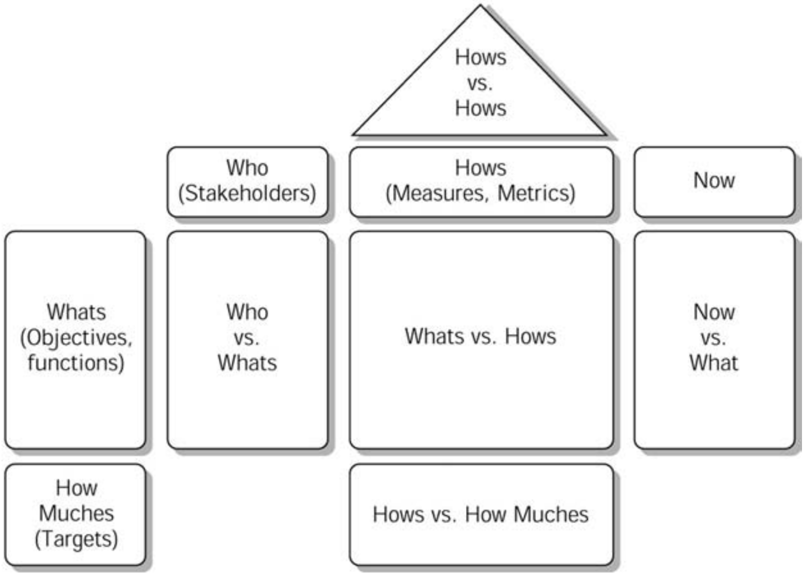
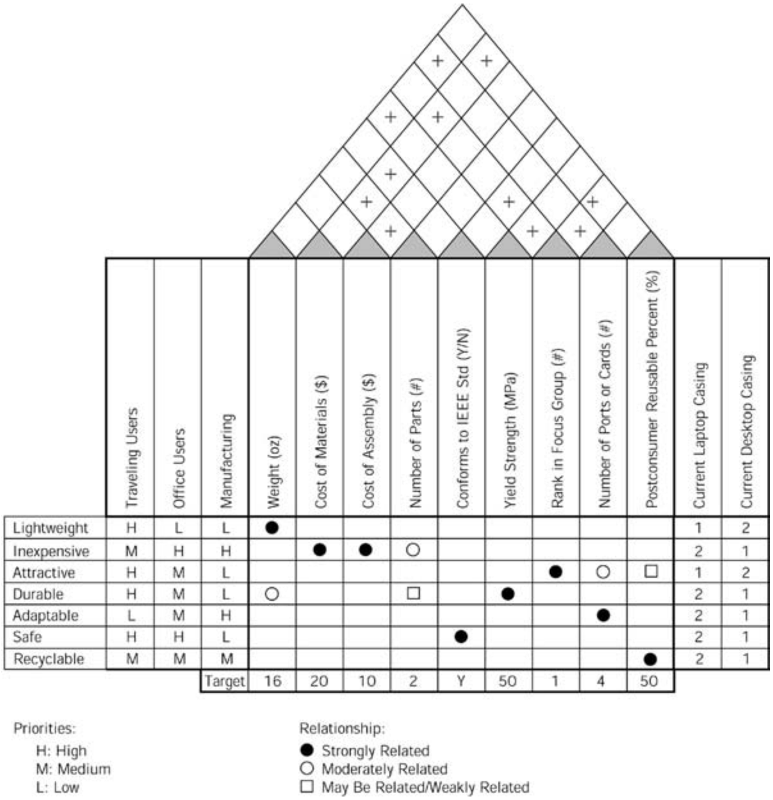



**Figure 6.6** An elementary abstraction of a *House of Quality* that displays and relates stakeholder interests, design attributes, measures, targets, and current products. The HoQ helps designers explore relationships among them. After Ullman (1997).

Consider using a HoQ to explore a housing for a laptop computer (Figure 6.7). A computer maker might want to explore designing computer housings that work for both office computers and laptops. The stakeholders include traveling users, office users, and the manufacturer's production group. In the *Who* versus *Whats* section, we see that travelers place a high priority on objectives lightweight and durable, while office users are more interested in inexpensive and adaptable. Then, if we had two existing designs, one a standard laptop case and the other a standard desktop/tower casing, we would see in the *Whats* versus *Hows* section that the costs (of raw materials and of assembly) are strongly related to inexpensive, while number of parts relates to inexpensive only weakly. Similarly, the number of cards and ports that can be accepted is also modestly related to inexpensive because they require added assembly work or more parts. *Now* versus *What* is the result of benchmarking the two existing design choices: It highlights the possibility that a "universal" housing might be able to satisfy more users in total if it can address the shortcomings of either design. Finally, the roof of the house shows some of the relationships and trade-offs that designers will need to consider. Making the case lighter, for example, is likely to trade off negatively with resistance to forces. Similarly, increasing the number of parts is likely to result in higher assembly costs.

This simple example shows that the house of quality can help tie together many of the concepts we have thus far described. When would we introduce a HoQ into our design process? There's no obvious answer. A HoQ is useful for gathering and organizing information, and for fostering discussions within a design team and with stakeholders. On the other hand, building a HoQ entails a lot of time and effort. So whether, and when, to build a HoQ are decisions that can only be made by a design team in the context of its design problem.




| | Traveling Users | Office Users | Manufacturing | Weight (oz) | Cost of Materials ($) | Cost of Assembly ($) | Number of Parts (#) | Conforms to IEEE Std (Y/N) | Yield Strength (MPa) | Rank in Focus Group (#) | Number of Ports or Cards (#) | Postconsumer Reusable Percent (%) | Current Laptop Casing | Current Desktop Casing |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Lightweight | H | L | L | ● | | | | | | | | | 1 | 2 |
| Inexpensive | M | H | H | | ● | ● | ○ | | | | | | 2 | 1 |
| Attractive | H | M | L | | | | | | | ● | ○ | □ | 1 | 2 |
| Durable | H | M | L | ○ | | | □ | | ● | | | | 2 | 1 |
| Adaptable | L | M | H | | | | | | | | ● | | 2 | 1 |
| Safe | H | H | L | | | | | ● | | | | | 2 | 1 |
| Recyclable | M | M | M | | | | | | | | | ● | 2 | 1 |
| | | | Target | 16 | 20 | 10 | 2 | Y | 50 | 1 | 4 | 50 | | |

**Figure 6.7** A first-draft HoQ for the design of a computer housing for a laptop and an office computer. Note that different users may have different priorities, and the roof of the house helps identify trade-offs between various objectives and features and behaviors.

## 6.4 FUNCTIONS FOR THE DANBURY ARM SUPPORT

We now continue our case study of how the E4 design teams designed a support arm for a student with CP.

The support arm must perform a number of functions in order to enable her writing and drawing. Teams A and B both used enumeration to develop their lists of functions, which are shown in Tables 6.2 and 6.3. Notice that many of Team A's functions are passive, that is, they are expressed as "enable . . . " While this is not unusual as a starting point for functional analysis, it is much better to express functions with active verbs because they are easier to translate into performance specifications. For example, "Enable size adjustability" might mean "Adjust size" or "Adjust to a (specified) range of sizes," or both. In this case, a passive formulation is ambiguous. Similarly, "Enable adjustable orientation" can be more clearly stated as "Adjust orientation," or as "Adjust to a (specified) range of orientations," or both.

Some of the functions listed in Table 6.2 are vague, so it is hard to identify means to perform those functions or to specify their actual performance. For example, the function

**TABLE 6.2 Team A's list of functions for the Danbury CP support arm**

| |
|---|
| Attach to something secure |
| Stay secure on arm |
| Support Jessica's arm |
| Decrease the magnitude of exaggeration |
| Enable size adjustability |
| Enable adjustable resistance/support |
| Enable adjustable orientation |
| Resist damage due to mishandling |
| Resist environmentally induced damage |
| Prevent physical pain |
| Provide comfort |

"Resist environmental damage," could almost be an objective, and does not identify particular environmental threats. (Similar concerns can be raised about "Prevent physical pain" and "Provide comfort.") To be fair, the team did provide some explanation of their functions. For example, on environmental issues the team wrote: "The device will sustain minimal damage if exposed to water, air moisture, dirt, dust, or other environmental factors in order to maximize the lifetime of the device." This explanation does provide more detail, but still leaves open "other environmental factors"—and it also suggests another objective, namely, "to maximize the lifetime of the device." Further, neither the stated function nor its explanation provide any quantification of how much exposure to various damaging agents the device is expected to survive. We don't mean to be unduly critical of our students' work, especially because expressing functions requires some very fine balancing: We have to write functions in sufficiently general terms that don't imply specific solutions, but we also need enough specificity so that their meaning is unambiguous and can be translated into meaningful specifications.

Team B enumerated a sharper, more concise list of functions displayed in Figure 6.3. Most of the functions listed in Table 6.3 also appear in Table 6.2, and the function "Dampen motion," seems to have the same intent as Team A's "Decrease the magnitude of exaggeration"; it could also be included in their function "Enable adjustable resistance/support" in Table 6.2. Similarly, the three adjustment functions of Table 6.2 (on size, resistance, and orientation) might be viewed as three articulated subfunctions of a single top-level function, "Enable adjustability"—which is more properly cast in Table 6.3 as the seemingly more

**TABLE 6.3 Team B's list of functions for the Danbury CP support arm**

| |
|---|
| Attach to arm |
| Attach to stabilizing point |
| Dampen motion |
| Allow for range of motion |
| Provide comfort |
| Provide adjustability |

active "Provide adjustability." Likewise, the two functions "Prevent physical pain" and "Provide comfort" of Table 6.2 can also be viewed as more detailed expressions of the function "Provide comfort" of Table 6.3.

It ought to be noted that while "enable" and "provide" are verbs that qualify as literal parts of verb–noun pairings, they are generally *passive* verbs. Functions are better understood and translated into good specifications when they are expressed in *active* verbs. For example, a sharper version of "enable adjustability" is "adjust to user size." Similarly, "cushion forces" is a more crisp formulation of "provide comfort."

While neither team was specifically asked to produce a formal list of design specifications, Team A did produce specifications for the 11 functions; they are displayed in Table 6.4. While this set of specifications served Team A adequately for this project, they would clearly not be accepted as a formal set of specifications because they are not the rigorous engineering statements needed to specify, measure, and test the functional performance of a design. It is also clear from Table 6.4 (as well as the student team's actual final report) that the team viewed the specifications more with an intent to consider

**TABLE 6.4 The performance specifications for Team A's 11 functions as listed in Table 6.2. Note that the requirements are qualitative, rather than "hard" quantitative specifications**

| Functions | Performance Specifications |
|---|---|
| Attach to something secure | |
| Attach securely to arm | Must not permit user to remove arm without assistance or use of free arm |
| Support Jane's arm | Once user raises arm to a specific height above arm rest, this height should be maintained without dropping or requiring user to apply muscle tension unless user chooses to change |
| Decrease the magnitude of the exaggeration of Student X's arm movements | As the user attempts to move her arm from point A to point B, the distance she ends at point B should be smaller than the distance if she were not using the device |
| Enable adjustability of size | Component containing/supporting user arm should have two settings: (1) fit to arm (to not allow play), (2) allow a little play |
| Enable adjustability of resistance/support | Three resistance settings: (1) none, (2) resist sudden jerks, (3) completely resist movement |
| Enable adjustability of orientation | Two stationary positions: (1) arm rest (just above arm rest of wheelchair), (2) work position (above table, next to canvas) |
| Resist damage due to mishandling | Maintain structure and form during mounting, dismount, and transport |
| Resist environmentally induced damage | Performance is not sensitive to dust, water, or paint |
| Prevent physical pain | User should not receive bruises, cuts, or experience strain |
| Provide comfort | Components containing/supporting arm contain some form of cushion or soft padding |

whether a final design was acceptable, which is understandable. But it is important to remember that specifications are used *prospectively*, to state precisely what has to be done to satisfactorily achieve a function, for example, "a structure must support a weight *X*," as opposed to the retrospective "our structure supported weight *Y*." (Both teams produced morph charts with their function lists, and they will be presented in Chapter 7.)

## 6.5 NOTES

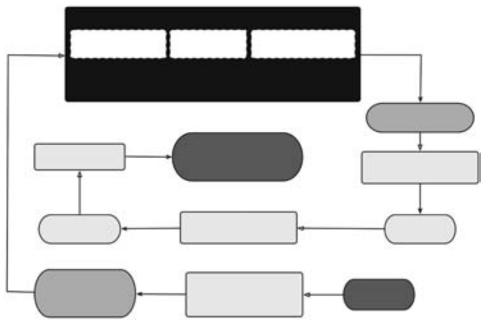
*Section 6.2*: Further details about engineering functions and requirements can be found in Ullman (1997). The term glass box method was coined in Jones (1992). The function–means tree used was developed by a former HMC colleague, James Rosenberg, to illustrate an example originally proposed in Akiyama (1991).

*Section 6.4*: The results from the Danbury arm support design project are from the final reports by Attarian et al. (2007) and Best et al. (2007).

CHAPTER 7

# CONCEPTUAL DESIGN: GENERATING DESIGN ALTERNATIVES

*How do I generate or create feasible designs?*



HAVING DEFINED a design problem by clarifying objectives, identifying constraints, and establishing functions, we now initiate its conceptual design by generating or creating design concepts. We will complete our discussion of conceptual design phase in Chapter 8 when we evaluate and choose among our design alternatives.

## 7.1 GENERATING THE "DESIGN SPACE," A SPACE OF ENGINEERING DESIGNS

How do we *generate* or *create* actual designs? We start by building a *design space*, an imaginary intellectual region of design alternatives that contains all of the potential solutions to our design problem. A design space is a useful notion that conveys a *feel* for the problem at hand: A *large design space* suggests a design domain with a large number of acceptable designs, or a design problem with a large number of design variables. While we can often look at a domain and intuit something about its design space (e.g., auto and building designs occupy very large design spaces), it is not clear how we identify a design space for unfamiliar or new devices. We now introduce the *morphological chart* as a formal tool for generating design spaces and for generating within those spaces a population of designs that perform the functions we specify. After that we will look at analogical thinking, another approach for generating design alternatives, and then offer one verbal and two graphical tools for coming up with designs in team-based activities.

### 7.1.1 Defining a Design Space by Generating a Morphological Chart

A *morphological chart* (aka a *morph chart*) is a matrix in which the leftmost column is a list of all of the principal functions that our design must perform and also some of the key features it must have. The list should be of a manageable size, and all of the entries should be at the same level of detail to help ensure consistency. Then, across from each of the functions or features, we list each of the different *means* of realizing the function or feature that we can think of. We strongly encourage *separating* functions from key features, for several reasons. First, we know that our design must be fully functional to satisfy our client's requirements. By putting all the functions together on the morph chart, we know we have addressed them all. The second reason we encourage separating functions from features is that a morph chart can quickly become rather large, and we may lose track of or confuse functions with key features. If we separate function from features at the outset, we can easily create two "design space" models in two separate charts, if necessary.

If we listed all the functions for the beverage container problem and arrayed the means corresponding to each to the right of each entry, we'd get the morph chart shown in Figure 7.1. We see that some functions have more means than others, for example, the function Contain Beverage has four means, while Resist Forces has only two. When we see a very small number of means this suggests that either we have a small design space (i.e., limited choices) or we have not fully explored the available design space.

We start building conceptual designs from the morph chart by noting that any feasible design must be *functionally complete*: every function, listed in the leftmost column must be achieved by that design. So we assemble designs by choosing one means from each row, and combine them into a functional design concept or scheme. Thus, we see in Figure 7.2a that one feasible design for the new juice container is a heat-sealed bag with a tear corner, thick walls, and a distinctive label, and another is a bottle with a twist top, made of a flexible material and with a distinctive shape.

| MEANS<br>FUNCTION | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Contain Liquid | Can | Bottle | Bag | Box | •••• | •••• |
| Fill and Seal Container | Fill and Heat Seal | Sealed Cap | Glue Container Material | Twist Top | Bottle Cap | |
| Empty Container | Pull Tab | Inserted Straw | Twist Top | Tear Corner | Unfold Container | Zipper |
| Resist Forces | Thick Walls | Flexible Materials | | | | |
| Identify Product | Shape of Container | Distinctive Label | Color | •••• | •••• | •••• |

**Figure 7.1** A morphological ("morph") chart for the juice container design problem with *functions* listed in the leftmost column. The *means* by which each can be implemented are arrayed along a row to each entry's right.

| MEANS / FUNCTIONS | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Contain Liquid | Can | Bottle | Bag | Box | •••• | •••• |
| Fill and Seal Container | Fill and Heat Seal | Sealed Cap | Glue Container Material | Twist Top | Bottle Cap | |
| Empty Container | Pull Tab | Inserted Straw | Twist Top | Tear Corner | Unfold Container | Zipper |
| Resist Forces | Thick Walls | Flexible Materials | | | | |
| Identify Product | Shape of Container | Distinctive Label | Color | •••• | •••• | •••• |

(a)

| MEANS / FUNCTIONS | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Contain Liquid | Can | Bottle | Bag | Box | •••• | •••• |
| Fill and Seal Container | Fill and Heat Seal | Sealed Cap | Glue Container Material | Twist Top | Bottle Cap | |
| Empty Container | Pull Tab | Inserted Straw | Twist Top | Tear Corner | Unfold Container | Zipper |
| Resist Forces | Thick Walls | Flexible Materials | | | | |
| Identify Product | Shape of Container | Distinctive Label | Color | •••• | •••• | •••• |

(b)

**Figure 7.2** The morphological ("morph") chart for the juice container design problem (Figure 7.1) is used to show (a) two feasible design alternatives whose means are dark and light shaded, and (b) two infeasible combinations whose means are also dark and light shaded.

The design generation method we have just described makes the morph chart into a spreadsheet with which we can "calculate" the number of potential designs. How many potential designs are there in that morph chart, that is, just how big is our design space? The answer reflects the *combinatorics* that result from combining a single means in a given row with each of the remaining means in all of the other rows. Thus, for the beverage container morph chart of Figure 7.1, the number of design alternatives could be as large as $4 \times 5 \times 6 \times 2 \times 3 = 720$.

While it seems that the design space for this simple example has suddenly become very large, it is important to recognize that not all of the combinations allowed by our combinatorial arithmetic are valid combinations, that is, not all of these 720 combinations are feasible designs. For example, we can see (Figure 7.2b) that we can't really design a bag with a zipper or a can with an unfolding corner!

Thus, our morphological chart provides both a tool to develop a design space and create design alternatives, and it provides an approach to *prune* that design space by identifying and excluding infeasible, incompatible alternatives. We exclude infeasible alternatives by, again, applying interface constraints, as well as physical principles and plain common sense.

We can use the morph chart to include key features as well as functions. In the juice container, for example, we might include a set of entries related to the materials we want to use, in which case we could distinguish between glass, plastic, Mylar, and cardboard. These features can help us understand our design space and conceptualize alternative designs, as well as generating more infeasible designs such as a glass bag.

There is a lot we can learn from our morph chart. Consider our problem with not having many means for Resisting Forces. The heart of this problem is that we need to consider resisting forces in more detail, distinguishing between Resisting Temperature and Resisting Shocks. At the same time, this may take us more deeply into particular designs than is appropriate at the conceptual stage. Once we have selected a concept, such as a bottle, we can increase its resistance to forces by thickening the walls, adopting appropriate structural elements, or wrapping it in a protective plastic. This shows us that it is important that we list functions (and features) at the same level of detail when we build a morph chart. Otherwise, we will find ourselves developing highly detailed designs at the conceptual stage, or still creating concepts even after we have settled on a scheme. Similarly, when doing a complex design task (e.g., designing a building), we don't want to worry about means for identifying exits or for opening doors while developing different concepts for moving between floors (e.g., elevators, escalators, and stairways).

We can also use morph charts to expand the design space for large, complex systems by listing principal subsystems in a starting column and then identifying various means of implementing each of those subsystems. For example, if we were designing a vehicle, we would need a subsystem Provide Power, which would have corresponding means like Gasoline, Diesel, Battery, Steam, and LNG. Each of these power sources is itself a subsystem that needs further detailed design, but we see we can use the morph chart idea to develop an array of subsystems to expand our range of design choices for a complex design. We might even choose to create a morph chart for some of these subsystems to help us appreciate the design choices implicit in our concepts and schemes.

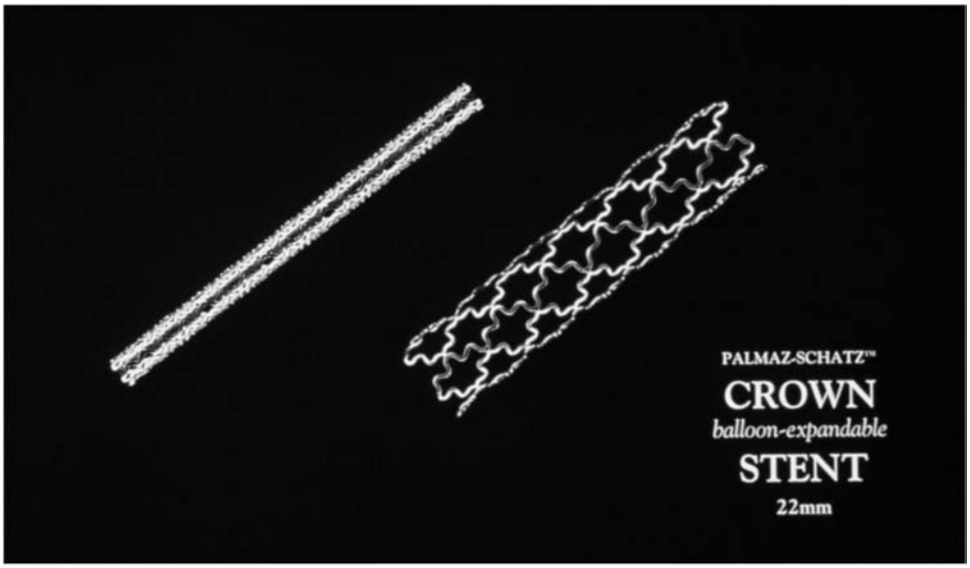
### 7.1.2 Thinking Metaphorically and Strategically

As we consider our morph chart, it is natural to ask where the ideas for the means and implementations come from. In this and the following sections, we look at a number of "ways of thinking" along with tools and techniques, all intended to help us come up with lots of creative and appropriate ideas. The first of these builds upon ways of thinking and speaking that we use in our everyday life, and shows us how they can be applied to engineering design.

A *metaphor* is a figure of speech that is used to give depth or color to the description of an object or process by likening it to another, usually more familiar, object or process. For example, when we describe engineering education as drinking from a fire hose, we mean to suggest that engineering students are exposed to a lot of knowledge quickly and under great pressure. We use metaphors to point out *analogies* between two different situations, that is, to suggest that there are parallels or similarities in the two sets of circumstances. Analogies can be very powerful tools in engineering design as illustrated by perhaps the most often cited: the Velcro fastener was designed by someone who drew a *direct analogy* between plant burrs that seem to stick to everything on which they're blown and the connecting fibers of the fastener.

We might use *symbolic analogies*, as when we "plant" ideas or talk about objectives "trees," because we are clearly drawing connections through some underlying symbolism. We can also stray into the realm of *fantasy analogies* by imagining something that is literally fantastic or beyond belief.

Fantasy analogies suggest another approach, sometimes called "thinking outside of the box." We are not very far past the time when many of the technologies we take for granted were thought to be outrageous ideas that were beyond belief. When Jules Verne published his classic *20,000 Leagues Under the Sea* in 1871, the idea of ships that could "sail" deep in the ocean was viewed as preposterous. Now, of course, submarines and seeing unfamiliar yet exciting life forms underwater are part of everyday experience. We cannot escape the idea that design teams might imagine the most outrageous solutions to a design problem and then seek ways to make such solutions possible. For example, airplanes that are invisible to radar were once considered far-fetched. The arterial stents used in angioplastic surgery (Figure 7.3) are also devices once thought to be impossible. Who would have believed that an engineering structure could be erected within the narrow confines of a human artery? And nowadays, of course, with the advent of nanotechnology, people are designing machines and devices just for use in human blood vessels.




Courtesy of Johnson & Johnson

**Figure 7.3** This is a PALMAZ-SCHATZ™ balloon expandable coronary *stent*, that is, a device used to maintain arterial shape and size so as to allow uninhibited and natural blood flow. Note how this structure resembles the kind of scaffolding often seen around building renovation and construction projects. Photo courtesy of Cordis, a Johnson & Johnson Company.

The stent suggests still another aspect of analogical thinking, namely, looking for *similar solutions*. The stent is similar in both intent and function to the scaffolding erected to support walls in mines and tunnels as they are being built. Thus, the stent and the scaffold are *like ideas*.

We could invert this idea by looking for *contrasting solutions* in which the conditions are so different, so contrasting, that a transfer of solutions would seem totally implausible. Here we would be looking for *opposite ideas*. Fairly obvious contrasts would be between strong and weak, light and dark, hot and cold, high and low, and so on. One example of using an opposite idea occurs in guitar design. Most guitars have their tuning pegs arrayed at the end of the neck. In order to make a portable guitar, one clever designer chose to put the tuning pegs at the other end of the strings, at the bottom of the body, in order to save space and thus increase the guitar's portability.

In addition to finding similar and contrasting solutions, we recognize a third category. *Contiguous solutions* are developed by thinking of *adjoining* (or *adjacent*) *ideas* in which we take advantage of natural connections between ideas, concepts, and artifacts. For example, chairs prompt us to think of tables, tires prompt us to think of cars, and so on. Contiguous solutions are distinguished from similar solutions by their adjacency, that is, bolts are adjacent to nuts and are contiguous solutions, while bolts and rivets serve identical fastening functions and are thus similar solutions.

The kinds of metaphorical thinking we have just described are also related to a characterization of design in terms of two different kinds of thinking:

- We do *divergent thinking* when we try to remove limits or barriers, hoping to increase our store of design ideas and choices. Speaking metaphorically, we "think outside of the box" or "stretch the envelope" when we want to expand our space of design alternatives.
- *We do convergent thinking* when we try to narrow our design space to focus on the best alternative(s). Again speaking metaphorically, we want to "stay within our game" and "know where the boundary/goal is" so we can converge on a solution within known boundaries or limits.

Perhaps we can best sum up this dilemma of choosing between thinking styles with an adapted metaphor: *Think outside the box, but stay within the physics!*

### 7.1.3 The 6–3–5 Method

Because so much engineering design is done in team settings, it is useful to consider tools and techniques that are well suited to teams. We consider three of these team-based activities. As before, our intent is to generate a rich set of ideas that can help us explore the entire design space.

The first of these team-based design generation activities is known as the *6–3–5 method*. Its name came from having *six* team members seated around a table to participate in this idea generation "game," each of whom writes down *three* design ideas, briefly expressed in key words and phrases. The six individual lists are then circulated past each of the remaining team members in a sequence of *five* rotations of *written* (only) comment and annotation: verbal communication or cross talk is not allowed. Thus, each list makes a complete circuit around the table, and each member of the team is stimulated in turn by the increasingly annotated lists of the other team members. When all of the participants have commented on each of the lists, the team lists, discusses, evaluates, and records all of the design ideas that have resulted from a group enhancement of the individual team members' ideas in a common visualization medium (e.g., a blackboard or a projection screen).

We can generalize this method to the "$m - 3 - (m - 1)$" method by starting with $m$ team members and using $m - 1$ rotations to complete a cycle. However, the logistics of everlengthening lists written on increasingly crowded sheets of paper, and of providing tables that seat more than six, suggest that six may be a "natural" upper limit for this activity. (In an academic setting, we would prefer fewer than six—ideally no more than four—on a project team.)

### 7.1.4 The C-Sketch Method

The *C-sketch method* starts with a team seated around a table, with each member *sketching one design idea* on a piece of paper, and then proceeds as does the 6–3–5 method. Each sketch is circulated through the team in the same fashion as the lists of ideas in the 6–3–5 method, with all of the annotations or proposed design modifications being written or sketched on the initial concept sketches. Again, the only permissible communication is by pencil on paper, with discussion following only after a complete cycle of sketching and modifying (as in the 6–3–5 method) has been completed. Research suggests that the C-sketch method can become unwieldy with even five team members due to the crowding of annotations and modifications on a given sketch. However, the C-sketch method is very appealing in an area such as mechanical design because there is strongly suggestive evidence that sketching is a natural form of thinking in mechanical device design. Research has also shown that drawings and diagrams facilitate the grouping of relevant information (usually added in marginal notes), and they help people to better visualize the objects being discussed.

### 7.1.5 The Gallery Method

The *gallery method* is a third approach to getting team reactions to design idea sketches, although the communication cycles are done differently. In the gallery method, team members first develop their individual, initial ideas within some allotted time, after which all of the resulting sketches are posted on a corkboard or a conference room whiteboard. This set of sketches serves as the backdrop for an open, group discussion of *all* of the posted ideas. Questions are asked, critiques are offered, and suggestions are made. Then each participant returns to her or his drawing and suitably modifies or revises it, again within a specified period of time, with the goal of producing a second-generation idea. The gallery method is thus both iterative and progressive, and there is no way to predict just how many cycles of individual idea generation and group discussion should be held. Our only recourse would be to invoke the idea of a utility plot and apply the *law of diminishing returns*: We proceed until a consensus emerges within the group that one more cycle will not gain much (or any) new information, then we quit because have reached a saturation plateau.

Note that the C-sketch and gallery methods provide contexts for committing design thoughts to paper by sketching. In fact, design teams do a variety of sketching and drawing activities, for a variety of purposes, and using a variety of technologies. We will discuss this (and show examples) in Chapter 9 and Appendix II, but we emphasize that the C-sketch and gallery methods need only rudimentary sketches, so *a designer need not be an artist to be a visual thinker.*

### 7.1.6 Guiding Thoughts on Design Generation

It is worth remembering that design generation is an exciting, creative activity, but it is goal-directed creative activity: It is designed to serve a known purpose, not to search for one. The goal may be imposed externally, as is often the case in engineering design firms, or internally, as in the development of a new product in a garage. But, *there is a goal toward which that creative activity is aimed*.

It is also worth remembering that *creative activity requires work*. As Thomas Edison famously said, "Invention is 99% perspiration and 1% inspiration." In other words, we have to be willing to do some serious work if we expect to be successful at generating design alternatives. Therefore, in order to do good, goal-directed design generation, we ask: Beyond the morph chart, beyond the team-based tools, what else can we do to generate design ideas? Or, how can we usefully navigate, expand or, if needed, contract our design space?

## 7.2 NAVIGATING, EXPANDING, AND CONTRACTING DESIGN SPACES

We began our discussion of design generation by proposing the morph chart as a formal tool for identifying spaces of designs that are populated by individual design alternatives. As we gain design experience, we find it more natural to think about design spaces and classes (not just individual designs) because we see commonalities across types of products or devices. Further, our experiences will increasingly turn toward designing more complex engineering systems, meaning that we'll have to design more subsystems and components, and we'll have to combine and connect these subsidiary individual designs. Thus, we offer a few thoughts about thinking effectively at the design space level.

### 7.2.1 Navigating Design Spaces

Large design spaces are complex because of the combinatorial possibilities that emerge when hundreds or thousands of design variables must be assigned. Design spaces are also complex because of interactions between subsystems and components, even when the number of choices is not overwhelming. In fact, one aspect of design complexity is that collaboration with many specialists is often critical because it is rare that a single engineer knows enough to make all of the design choices and analyses.

Two designed objects that have large design spaces are passenger aircraft (e.g., the Boeing 747), and major office buildings (e.g., Chicago's Sears Tower). A 747 has six million different parts, and we can only imagine how many parts there are in a 100-story building, from window frames and structural rivets to water faucets and elevator buttons.

With so many parts, there are still more design variables and design choices. Yet, while both the 747 and the Sears Tower have very large design spaces, these devices differ from one another because their performances present different challenges and different constraints. Architect and structural designers of a skyscraper have far more choices for the shape, footprint, and structural configuration of a high-rise than do aeronautical engineers, who must design fuselages and wings within strict aerodynamic constraints. While a building's weight is important as its number of floors and occupants rise, and while high-rise buildings' shapes are analyzed and tested for their response to wind, they are subject to fewer constraints than are the payload and aerodynamic shape of aircraft.

A *small* or *bounded design space*, on the other hand, conveys the image of a design problem in which the number of potential designs is limited or small, or the number of design variables is small and they, in turn, take on values within limited ranges. Thus, the design of individual components of large systems often occurs within small design spaces. For example, the design of windows in both aircraft and buildings is so constrained by opening sizes and materials that their design spaces are relatively small. Similarly, the range of framing patterns for low-rise industrial warehouse buildings is limited, as are the kinds of structural members and connections used to make up those structural frames.

One of the complications of large design spaces stems from the fact that many design variables are highly dependent either on choices already made or on those yet to be made. We attack such complex design spaces by applying the idea of *decomposition*, or *divide and conquer*: divide (or break into pieces) a complex problem into subproblems that are more readily solved. Designs of airplanes, for example, can be decomposed into subproblems: the wings; the fuselage; the avionics; the tail; the galley; the passenger compartment; and so on. In other words, the overall design problem and space are broken into manageable pieces that are taken on one at a time. Indeed, the morphological chart is particularly suited to (1) decomposing the overall functionality of a design into its constituent subfunctions; (2) identifying means for achieving each of those subfunctions; and (3) supporting the *synthesis*, or composition and *re*composition, of feasible design solutions.

### 7.2.2 Expanding a Design Space When It Is Too Small

We sometimes feel that our design space is too small, that we may not have enough options. There are things we can do to expand our design space, and they fall into various kinds of information gathering, including:

- *Conduct literature reviews* to determine the state of the art and identify prior work in the field. This includes locating and studying previous solutions, product advertising, vendor literature, as well as handbooks, compendia of material properties, design and legal codes, and so on. For example, *The Thomas Register* lists more than one million manufacturers of the kinds of systems and components used in mechanical design. Further, much more material is becoming available on the World Wide Web, although it is risky to assume that *all* information is both web-available and technically correct.
- *Conduct a patent search* to identify available technologies, to avoid "reinventing the wheel" and to leverage our thinking by building on what we already know about a still-emerging design. Patents are a kind of *intellectual property*: Patent holders

identified by the U.S. Patent Office (USPTO) are credited for having invented new devices or discovered new ways to do things. The USPTO awards two kinds of patents: *design patents* on the *form* or appearance or "look and feel" of an idea; and *utility patents* for *functions*, that is, on how to do something or make something happen.

- *Benchmark* existing products to evaluate how *well* they perform.
- *Reverse engineer* devices to see *how* functions are performed and to identify alternate means of performing similar.

### 7.2.3 Contracting a Design Space When It Is Too Large

We often feel that our design space is too large, that we have too many options, so we need to *prune* or contract our space to make it more manageable. There are several pragmatic guideposts for narrowing a search space, including:

- Check for *external constraints* that affect the design. For example, make sure the design team's competence maps onto the design problem posed (e.g., it may be more comfortable designing tricycles than high-tech performance bikes). For another example, be sure that the manufacturing capabilities are available (e.g., a team should avoid designing a bike made of composite materials if the only available manufacturing facility forms and connects metals).

*Invoke and apply constraints*, in much the same way we did just above while assessing the presence of external constraints.

*Freeze the number of features and behaviors* being considered to avoid details that are unlikely to seriously affect the design at this point (e.g., the color of a bike or car is not worth noting in a design's early stages.)

*Impose some order on the list*, perhaps by harking back to data gathered during problem definition that suggests that particular functions or features are more important.

*"Get real!"* or, in other words, apply *common sense* to rule out infeasible ideas.

## 7.3 GENERATING DESIGNS FOR THE DANBURY ARM SUPPORT

We now return to following the two design teams working on the arm support for the CP-afflicted student at the Danbury School. Starting with the functions they had already identified, each team built a morph chart: Team A's, based on Table 6.2, is shown in Figure 7.4; and Team B's, based on Table 6.3, is displayed in Figure 7.5. In addition, although to different degrees, the teams did research on the availability of devices that were intended to serve the same functions for similar users. (It is worth noting, too, that teams in HMC's E4 design experience are routinely told that it is perfectly acceptable to recommend that the client buy an existing product ***provided*** they can identify an existing product that meets the client's objectives, performs the specified functions, and satisfies the problem's constraints. The teams are told this in part to let them know that this is a

| Functions | Means | | | | | |
|---|---|---|---|---|---|---|
| Attach to something secure | Clamp to chair armrest | Strap to chair backrest | Clamp to table | Strap to user's arm using velcro | | |
| Attach securely to arm | Velcro straps | Buckles | Clamps | Sleeves | Ties | Zippers |
| Support Jane's arm | Stationary support cup attached to an arm under student's elbow | Mobile support cup attached to an arm under student's elbow | Support frame with sliding bar | Free-swinging sling | Brace | Pulley system |
| Decrease the magnitude of the exaggeration of Jane's arm movements | Dashpots | Torsion springs | Elastic wires/ cables | Brake pads on hinges | | |
| Enable adjustability of size | Adjustable straps | "Baseball cap" snaps | Elastic material | Telescoping extensions | Variable elastic strap length | Clipped elastic straps |
| Enable adjustability of resistance/ support | Tightening screw | Adjustable brake pads | Pads to compress torsion springs | Reel to shorten elastic wires/ cables | Torsion springs | Variable air resistance |
| Enable adjustability of orientation | Telescoping rods | Lockable hinged arm | Sliding rail | Lockable pivot disks | Braking pivot disks | Casters for pivot disks |
| Resist damage due to mishandling | Slip covers | Rubber padding | Foam padding | Nondeformable materials | Nonbreakable materials | |
| Resist environmentally induced damage | Covers over parts with small crevices | Waterproof material | Rustproof material | | | |
| Prevent physical pain | Emergency release | Cover moving parts | Cover sharp edges | | | |
| Provide comfort | Soft coverings | Soft padding | Air holes | Breathable material | | |

**Figure 7.4** Excerpts from Team A's morph chart for the Danbury CP arm support. Adapted from Attarian et al. (2007).

| Functions | Means | | | | | |
|---|---|---|---|---|---|---|
| Attach to arm | Air pressure cuff | Elbow pad | Velcro straps | Rope, string | Arm rest | Sleeve, rings |
| Attach to stabilizing point | Clamps (one or two) | Nut and bolt | Harness | Magnets | Chemistry-style Clamps | |
| Dampen motion | Pneumatics | Flywheel | Seatbelt | Elastics | Viscous fluid | Friction |
| Allow for range of motion | Telescopic rods | Rails | | | | |
| Provide comfort | Pillows | Air cushion | Foam padding | Gel cushioning | | |
| Provide adjustability | Jack | Nuts and bolts | Pressure | Slide rails | Piston | |

**Figure 7.5** Team B's morph chart for the Danbury CP arm support. Adapted from Best et al. (2007).

legitimate outcome, and in part to encourage the teams to do their research so as to avoid "reinventing the wheel"!) A comparison of the two morph charts reinforces our earlier observations about focused thinking because Team B's is a more bounded morph chart (and design space) reflecting a sharper, more concise list of functions. Similarly, the three passive functions ("Enable adjustability of . . . ") shown in Team A's chart, and their associated means, might once again be viewed as three simple articulations of subfunctions and means of the top-level function, "Enable adjustability." In particular, if the overall objective is "Provide adjustability," is this the right time in the design process to consider details of different kinds of adjustments?

We also note that both of these morph charts are very large, that is, they have very large numbers of possible combinations: 13,310 (i.e., $11 \times 11 \times 11 \times 10$) alternatives for the (partial) chart of Figure 7.4 and 7200 for the smaller morph chart of Figure 7.5. These are overwhelmingly large numbers of outcomes for this design problem, which suggests the teams should think strategically about grouping and reorganizing the functions and the resultant design alternatives.

Both teams followed similar approaches: they blended possibilities stemming from their morph charts with information gained from their research and with their own experience-based judgments and gut-level feelings. Team A developed three designs based on their morph chart, as can be seen in the marked-up version of their morph chart in Figure 7.6. They called those three designs: "Sling"; "Sliding Bars"; and "Support Arm." We can also see from Figure 7.4 that there was a fair amount of overlap among their three designs, with several different means being used in more than one design. That may not be terribly surprising, given the nature of this particular design problem.

Team B took a less structured approach in which they decomposed the overall design into: a "structure component" that connected Jessica's arm to a structure that would maximize her range of motion and provide comfort and adjustability; and a "dampening component" that would minimize her involuntary contractions without restricting her

| Functions | Means | | | | | |
|---|---|---|---|---|---|---|
| Attach to something secure | Clamp to chair armrest **1,3** | Strap to chair backrest | Clamp to table **2** | Strap to user's arm using velcro | | |
| Attach securely to arm | Velcro straps **2,3** | Buckles | Clamps | Sleeves **1** | Ties | Zippers |
| Support Jane's arm | Stationary support cup attached to an arm under student's elbow | Mobile support cup attached to an arm under student's elbow **2,3** | Support frame with sliding bar | Free-swinging sling **1** | Brace | Pulley system |
| Decrease the magnitude of the exaggeration of Jane's arm movements | Dashpots **2,3** | Torsion springs **1** | Elastic wires/ cables **1** | Brake pads on hinges | | |
| Enable adjustability of size | Adjustable syrups **2,3** | "Baseball cap" snaps | Elastic material | Telescoping extensions **1,3** | Variable elastic strap length **1** | Clipped elastic straps |
| Enable adjustability of resistance/ support | Tightening screw **2** | Adjustable brake pads | Pads to compress torsion springs | Reel to shorten elastic wires/ cables **1** | Torsion springs **1** | Variable air resistance **2,3** |
| Enable adjustability of orientation | Telescoping rods **1,3** | Lockable hinged arm **3** | Sliding rail **2** | Lockable pivot disks | Braking pivot disks | Casters for pivot disks |
| Resist damage due to mishandling | Slip covers | Rubber padding | Foam padding | Nondeformable materials **1** | Nonbreakable materials **2,3** | |
| Resist environmentally induced damage | Covers over parts with small crevices **2** | Waterproof material **1,2,3** | Rustproof material | | | |
| Prevent physical pain | Emergency release | Cover moving parts **1,2,3** | Cover sharp edges **1,2,3** | | | |
| Provide comfort | Soft coverings | Soft padding **2,3** | Air holes | Breathable material **2** | | |

**Figure 7.6** A marked-up copy of the excerpts of Team A's morph chart for the Danbury CP arm support that shows how their three designs were assembled from their morph chart (Figure 7.4). Note that Team A often used two means to achieve a given function. The three designs are: (1) "Sling"; (2) "Sliding Bars"; and (3) "Support Arm." Adapted from Attarian et al. (2007).

voluntary motion or range of motion. The three designs were identified by Team B as "Dually-Hinged Structure with Rail," "Dually Hinged Structure," and "Ball and Socket Structure."

Some of the design sketches and drawings for the design concepts produced by the teams are shown in Figures 11.1 and 11.3 as part of our discussion of sketching, drawing, and prototyping. But is interesting to look ahead and see some of the most immediately tangible fruits (and joys) of successful conceptual design. Teams A and B selected the particular designs shown in Figures 11.1 and 11.3, respectively, after they applied their particular metrics to their corresponding objectives (see Figure 4.3 and Table 4.10).

## 7.4 NOTES

*Section 7.1*: Zwicki (1948) originated the idea of a morphological chart. Further discussion and examples of morph charts can be found in Cross (1994), Jones (1992), and Hubke (1988).

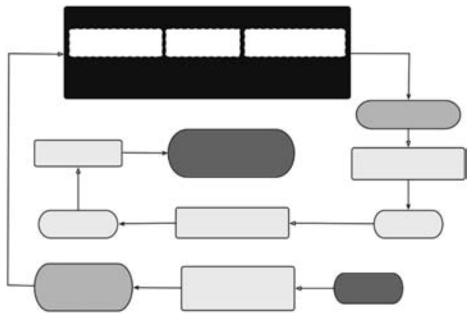
*Section 7.2*: The address of the USPO's website is www.uspto.gov. Another often-used website is www.ibm.com/patents. Group methods of idea generation are explored and described in Shah (1998). Approaches to creativity and analogical thinking in a group setting are described in Hays (1992).

*Section 7.3*: The results for the Danbury arm support design project are taken from final reports by Attarian et al. (2007) and Best et al. (2007).

CHAPTER 8

# CONCEPTUAL DESIGN: EVALUATING DESIGN ALTERNATIVES AND CHOOSING A DESIGN

*Which design should I choose? Which design is "best"?*



NOW WE will finish our work on conceptual design by evaluating design concepts to see which ones are worth pursuing. Because choices among alternative designs invariably involve subjective judgments, perhaps the most important lesson to be learned here is that we must always take care not to mistake things to which we somehow assign numbers for those things we can truly measure.

## 8.1 APPLYING METRICS TO OBJECTIVES: SELECTING THE PREFERRED DESIGN

With several feasible designs in hand, we now turn to choosing a "best" or preferred design—and choose we must because rarely are the resources (e.g., time, money, and personnel) available to fully develop more than one design scheme, never mind all of our alternatives. Whether we found our alternatives using a morph chart or a less structured approach, we have to "pick a winner" from among the identified options for further elaboration, testing, and evaluation. And how do we pick that winner? What makes the most sense is to choose the design (or designs) that best meet the client's objectives. Recall that we introduced metrics as a way to measure the achievement of objectives. Now we put those metrics to use to assess how close we've come to meeting each of the client's objectives, and combine our insights from the measurements for the individual objectives to reflect some sense of how the client's overall set of objectives are achieved.

Having said that, there are three potential problems in this process. First, we really need to try and limit our analysis to the client's *most important objectives*. This reflects common sense, but it also reminds us to avoid drowning useful information in a sea of relatively unimportant data. Second, we have to remember to establish our metrics with common sense of scale, so as not to mistakenly over- or underemphasize some results. Third, and last, when examining metrics results for several different objectives, we have to keep in mind that our information necessarily reflects a fair amount of subjectivity. This is especially the case when we rank objectives in terms of order of relative importance, and even in the application of metrics because many of those metrics reflect *qualitative*—not measurable, quantitative—results. At the most fundamental level, the results of our metrics should be thought of more as indicating a clear sense of direction than an algorithm or numerical solution.

We will discuss three methods for choosing from among a set of alternative designs or concepts: the numerical evaluation matrix; the priority checkmark method; and the best-of-class chart. The three selection methods explicitly link design alternatives to *ordered unweighted design objectives*. Ordered objectives cannot be scaled on a mathematically meaningful ruler, so we must be very cautious as we strive to bring order to judgments and assessments that are subjective at their root. Just as professors give grades to encapsulate judgments about how well students have mastered concepts, ideas, and methods, designers try to integrate the best judgments in a sensible and orderly manner. And we must always use common sense when we look at the results of applying any such method.

No matter what kind of chart or other decision support technique we apply, our first step should always be to check that each alternative satisfies all of the applicable constraints: *design alternatives that don't meet constraints must be immediately rejected as infeasible*. Having said that, as we describe our three selection methods, we will show how the applicable constraints can be applied to narrow the design space accordingly.

### 8.1.1 Numerical Evaluation Matrices

To examine the numerical evaluation matrix, we can revisit the juice container problem. Consider further that our design process has led us to four alternatives, each of which can be seen as a feasible design in the morph chart in Figure 7.1: a glass bottle with a distinctive shape, an aluminum can with a pull tab and a clever label, a Mylar bag where the juice is accessed via a straw, and a polyethylene bottle with a screw cap. We show in Table 8.1 a *numerical evaluation matrix* for this situation. This matrix shows both constraints (upper rows) and objectives (lower rows) in the left-hand column. (For simplicity's sake, we have limited our objectives to a subset of ones the client has indicated as particularly important.) We can immediately rule out glass bottles and aluminum containers because they violate a constraint because of their potential for sharp edges. Thus, we need to evaluate only two designs, a Mylar bag and a polyethylene bottle, against the metrics for the relevant objectives, which are environmentally benign, easy to distribute, and long shelf life. In general, we would try to limit the number of decisive objectives to the top two or three because it is difficult to mediate among more than two or three objectives at one time. If we

**TABLE 8.1 A *numerical evaluation matrix* for the juice container design problem. Note that only three of the six objectives originally identified for this design are utilized here, in part because we think these three objectives are more important than the other three, and in part because we have metrics (and presumably data) for these three objectives**

| Design<br>Constraints (C) and Objectives (O) | Glass Bottle, with Twist-Off Cap | Aluminum Can, with Pull-Tab | Polyethylene Bottle, with Twist-Off Cap | Mylar Bag, with Straw |
|---|---|---|---|---|
| C: No sharp edges | × | × | | |
| C: Chemically inert | | | | |
| O: Environmentally benign | | | 80 | 40 |
| O: Easy to distribute | | | 40 | 60 |
| O: Long shelf life | | | 90 | 100 |

apply the corresponding metrics, we get the results shown in the two rightmost columns of Table 8.1.

It is tempting to sum the data from these columns, but we should not. There is simply no basis for doing so, since it would imply that a score in one metric is somehow translatable into another. Indeed, adding the scores would suggest that the metric results could be considered as equally rank ordered, which we learned is not the case when we examined pairwise comparison charts. So what do the numbers mean? First of all, they allow us to see if one design is what as known as Pareto optimal, that is, superior in one or more dimensions, and at least equal in all the others. In this case, neither design is Pareto optimal. One design, the polyethylene bottle seems strongly superior with respect to the environmental objective, inferior with respect to distribution issues, and essentially tied in terms of shelf life. (Had we summed up the numbers, the particular ways in which one design is superior or inferior might well be masked, in addition to the problems mentioned above.) These values can be used to work with the client (and perhaps users) to revisit the objectives. It is not uncommon for a client to change their mind about relative rankings in order to get a very strong winner in some other dimension, especially if the design process has taken a long time since the initial consideration. There is more to be said about the individual component values that the different designs achieve.

This marked contrast between the designs with respect to different objectives suggests that designers might wind up choosing different designs *based on the values of their clients*. For example, one client might value an environmentally friendly container above all else, to the point of making the objective *environmentally benign* its dominant or even sole objective. That client would choose the polyethylene bottle. Another client might value ease of distribution above all else, perhaps because its principal concern is to get its new juice distributed into markets in poor countries as quickly as possible. With this value in mind, they might choose the Mylar bag with the same degree of rationality as NBC choosing the polyethylene bottle. (We can see this in the design decisions surrounding food packaging for disaster relief, where environmentalism is considered important, but secondary to the ability to get help where it is most needed.)

The results in Table 8.1 give the metric results awarded for each objective of each design, that is, they reflect solely the application of the metrics to the two design alternatives. They are independent of the client. In fact, there's nothing in Table 8.1 that identifies the client, which is as it should be. If our metrics produced different values for different clients (for the same product), we would have to wonder whether there was a defect in the testing process or in how the test results were reported. Ideally, a design team's metrics and associated testing procedures should not change with whoever is applying or making measurements for the metrics.

### 8.1.2 Priority Checkmark Method

The *priority checkmark method* is a simpler, qualitative version of the numerical evaluation matrix we have just described. We simply rank the objectives as high, medium, or low in priority. Objectives with high priority are given three checks, those with medium priority are given two checks, while objectives with low priority are given only one check, as shown in Table 8.2 for the same designs of Table 8.1. Similarly, metric results are assigned as 1 if they are awarded more than some arbitrary, but high value, such as 70 points (on a scale of 0–100), and as 0 if their award is less than the target value. Thus, a design alternative that meets an objective in a "satisfactory" way is then marked with one or more checks, as shown in Table 8.2. This method is easy to use, makes the setting of priorities rather simple, and is readily understood by clients and by other parties. On the other hand, the priority checkmark loses considerable information that may be useful in differentiating between relatively close alternatives. The actual metric results, for example, are no longer presented for consideration and discussion. In addition, because

**TABLE 8.2 A *priority benchmark chart* for the juice container design problem. This chart qualitatively reflects a client's values in terms of the priority assigned to each objective, so it uses the ordering in the PCC of Figure 4.4**

| Design Constraints and Objectives | Priority (✓) | Glass Bottle, with Twist-Off Cap | Aluminum Can, with Pull-Tab | Polyethylene Bottle, with Twist-Off Cap | Mylar Bag, with Straw |
|---|---|---|---|---|---|
| C: No sharp edges | | × | × | | |
| C: Chemically inert | | | | | |
| O: Environmentally benign | ✓✓✓ | | | 1 × ✓✓✓<br>✓✓✓ | 0 × ✓✓✓<br>●●●● |
| O: Easy to distribute | ✓ | | | 0 × ✓<br>●●●● | 1 × ✓<br>✓ |
| O: Long Shelf Life | ✓✓ | | | 1 × ✓✓<br>✓✓ | 1 × ✓✓<br>✓✓ |

of its binary nature, this method may lead to results that appear to be more disparate than they really are. Designers need to be particularly careful not to succumb to the temptation to "cook the results" in choosing the thresholds.

### 8.1.3 The Best-of-Class Chart

Our last method for ranking alternatives is the *best-of-class chart*. For each objective, we assign scores to each design alternative that start from 1 for the alternative that meets that objective best, increasing to 2 for second-best, and so on, until the alternative that met the objective worst is given a score equal to the number of alternatives being considered. So if there are seven alternatives, then the best at meeting a particular objective would receive a 1, and the worst a 7. Ties are allowed (e.g., two alternatives are considered "best" and so are tied for first) and are handled by splitting the available rankings (e.g., two "firsts" would each get a score of $(1+2)/2 = 1.5$, and a tie between the "second" and "third" would get $(2+3)/2 = 2.5$). Once again, we do not sum the results, since this would imply that all the objectives being considered have equal weight. Rather, the scores allow us to see if we have a design that is Pareto optimal (best in all categories), or at least best in the most important (i.e., highest ranked) objectives, and to discuss what the various high and low performances will mean for the final concept or scheme adopted.

Table 8.3 shows a best-of-class chart for the juice container example. Notice that this approach allows the designer and client to see how each design ranked with respect to the objectives, but gives no information on the actual scoring. If, for example, the Mylar bag was very close on the environmental scores, we might choose to treat that as a near tie, and select it over the alternative.

The best-of-class approach has advantages and disadvantages. One advantage is that it allows us to evaluate alternatives with respect to the results for each metric, rather than simply treat as a binary yes/no decision, as we did with priority checkmarks. It is also relatively easy to implement and explain. The ranking methods also allow for qualitative evaluations and judgments, since the persons applying the metric do not

**TABLE 8.3 A *best-of-class chart* for the juice container design problem. This chart presents the rank ordering of the metrics results for each acceptable design. Notice that in this case, the client and the designer will need to select between the winner for the highest objective, or a design that wins on both of the other ones**

| Design<br>Constraints (C) and Objectives (O) | Glass Bottle, with Twist-Off Cap | Aluminum Can, with Pull-Tab | Polyethylene Bottle, with Twist-Off Cap | Mylar Bag, with Straw |
|---|---|---|---|---|
| C: No sharp edges | * | * | | |
| C: Chemically inert | | | | |
| O: Environmentally benign | | | 1 | 2 |
| O: Easy to distribute | | | 2 | 1 |
| O: Long shelf life | | | 2 | 1 |

have to have a formal quantitative test for each objective. (This can be particularly helpful when the concepts are too expensive, difficult, or time-consuming to produce as prototypes.) The method can also be done by individual team members or by a design team as a whole to make explicit any differences in rankings or approaches. The disadvantages of the best-of-class approach are that it encourages evaluation based on opinion rather than testing or actual metrics, and it may lead to a moral hazard akin to that attached to priority checkmarks, that is, the temptation to fudge the results or cook the books. It also shows only the rankings, but not the actual score. As we saw in our juice container example, we do not know if the first and second results are close or not, which could be important information. Best of class can be particularly helpful if there are many alternatives and we want to narrow or consultative and thoughtful process to the top few.

### 8.1.4 An Important Reminder About Design Evaluation

No matter which of the three selection methods is used, design evaluation and selection demand careful, thoughtful judgment. First and foremost, as we have cautioned earlier, the ordinal rankings of the objectives obtained using PCCs *cannot* be meaningfully scaled or weighted. To draw a crude analogy, think of being at the finishing line of a race without a clock: we can observe the order in which racers finish, but we cannot measure how fast (i.e., how well) they finish the course. Similarly, while we can measure *ranking* with a PCC, we cannot measure or scale objectives' weights from their PCC order of finish: *A PCC's ordinal rankings cannot be weighted or scaled*. This means that we also cannot simply sum our results, since that would imply that all our objectives are being weighted with an equal value of 1.

Further, *we must always exercise common sense as we are evaluating results*. If the metrics results for two alternative designs are relatively close, they should be treated as effectively equal, unless there are other unevaluated strengths or weaknesses. Further, if we are surprised by our evaluations, we should ask whether our expectations were wrong, our measurements were consistently applied, or whether our rankings and our metrics are appropriate to the problem.

Still further, if the results do meet our expectations, we should ask whether we have done our evaluation fairly, or we have just reinforced some biases or preconceived ideas. Finally, it might be wise to check whether the constraints used to eliminate designs are truly binding.

In brief, *there is no excuse for accepting results blindly and uncritically.*

## 8.2 EVALUATING DESIGNS FOR THE DANBURY ARM SUPPORT

We now return to following the two design teams working on the arm support for the CP-afflicted student at the Danbury School. We noted in Chapter 7 that Team B developed three designs that we identified earlier as "Dually-Hinged Structure with Rail," "Dually-Hinged Structure," and "Ball and Socket Structure." They applied their metrics (previously given in Table 4.11) as shown here in Table 8.4. Team B's report stated that, "As

**TABLE 8.4 This table shows the results of applying Team B's metrics to their three designs**

| Objective | Dually-Hinged Structure with Rail | Dually-Hinged Structure | Rail and Socket Structure |
|---|---|---|---|
| Safety/10 | 8 | 9 | 9 |
| Stabilization/10 | 5 | 6 | 4 |
| Comfort/10 | 8 | 6 | 8 |
| Nonrestrictive/10 | 9 | 8 | 9 |
| Ease of Installation/10 | 8 | 8 | 6 |
| Durability/10 | 9 | 10 | 7 |
| Adjustability/10 | 8 | 6 | 8 |
| Low Cost/10 | 7 | 8 | 8 |

Adapted from Best et al. (2007).

indicated by [the] applied metrics, the dually-hinged structure incorporating a rail overcame competing designs." Their results highlight that designers cannot simply rely on numerical outcomes blindly. In this case, the selected design scored as high or higher on their top two objectives, but had mixed results on some of the others. This is often the case—designers must ultimately exercise informed judgment in consultation with their client. The dually hinged structure was declared the winner and, with some further modifications, was prototyped.

Team A developed three designs from their morph chart (see Figure 7.6). They called those three designs "Sling," "Sliding Bars," and "Support Arm." According to team A's final report,

> *"The evaluation of design alternatives consisted of (1) assigning percent weights to each objective, and (2) assigning a score to each design alternative for each objective, by estimating how well a design met an objective using metrics. After tallying the scores for the three devices, the support arm had the highest score. Thus, the support arm was selected as the final design to be refined and constructed."*

Team A did not present any of the results they said they obtained, save for the qualitative evaluation of their final design that is shown in Table 8.5, which show evaluation "Conclusions" and "Results" for the metrics we had previously shown in Table 4.10. The team clearly did several things *it should not have*, including: weighting the objectives, not limiting the number of objectives used in this design evaluation, not indicating in any way the margins (absolute or even relative) of the summed scores for their three designs, and just generally not providing the underlying basis for their design decisions to the client (or to their faculty advisors!). They were fortunate that they were working in a relatively narrow design space, as we noted before, and that they had used their ingenuity and insights into currently available products to produce what seemed like a reasonable design.

**TABLE 8.5 Partial evaluation table of some of the most important objectives for the Danbury CP arm, as viewed by Team A, along with metrics, "Conclusion(s)," and final results**

| Objectives | Metrics | Conclusion | Result |
|---|---|---|---|
| 1. Minimize number of sharp edges | Number of sharp edges | Inherent sharp metal edges | Fail |
| 2. Minimize pinching | Number of pinching possibilities | User quite comfortable | Pass |
| 3. Finger friendly | Number of places on device to get finger caught | Not safe to handle | Fail |
| 4. Durable | Disconfiguration, misalignment of device after regular use | Insecure mount, misaligns | Fail |
| 5. Remain secure on user | Conditions under which device remains securely attached to user | Arm remains attached to device | Pass |
| 6. Maintain stable position | Conditions where position, orientation of device maintain mounting setting | Insecure mount | Fail |
| 7. Minimize cost | Estimate dollar amount | Less than other products | Pass |
| 8. Normalize arm movement | User ability to draw straight line compared to ability to do so without the device | Failure in extensibility hurts use | Fail |
| 9. Maximize range of voluntary motion | Degree of freedom in motion of wrist, elbow, arm, and torso | Comfortable range of motion, except torso forward bend | Pass (except torso) |
| 10. Movable while in use | Required assembly condition to move device | Does not require disassembly | Pass |
| 11. Transportable | Necessary level of disassembly for movement | Does not require disassembly | Pass |
| 12. Usable by multiple students | Range of permissible arm sizes | Adjustable size permits various arm sizes | Pass |

Adapted from Attarian et al. (2007).

## 8.3 NOTES

*Section 8.1*: The three design evaluation methods discussed derive from Pugh's concept selection method, which is discussed in Pugh (1990), Ullman (1992, 1997), and Ulrich and Eppinger (1997, 2000).

*Section 8.2*: The results for the Danbury arm support design project are taken from Attarian et al. (2007) and Best et al. (2007), but have been modified and corrected for pedagogical reasons. Both teams attached weights to their metrics and summed them, which is incorrect. Rather than present two incorrect examples, the authors elected to present one correct, that is, unweighted, example.

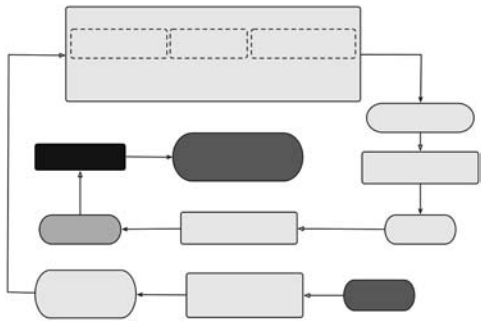
PART III

# DESIGN COMMUNICATION

CHAPTER 9

# COMMUNICATING DESIGNS GRAPHICALLY

*Here's my design; can you make it?*



BEING ABLE to communicate effectively is a critical skill for engineers. We communicate in oral presentations, and through written documents, and technical drawings. We also communicate individually and as members of design teams. We communicate with our client: when we define the design problem; while we work through the design process; and when we portray our final design in standardized, detailed drawings so it can be built. We communicate when we build models or prototypes to demonstrate or evaluate our design's effectiveness. And perhaps as important as anything else, we also communicate when we take our ideas from our heads and commit them to paper. We devote this chapter to creating design drawings, an essential tool for effective engineering communication.

## 9.1 ENGINEERING SKETCHES AND DRAWINGS SPEAK TO MANY AUDIENCES

Drawing is very important in design because a lot of information is created and transmitted in the drawing process. Design drawings include sketches, freehand drawings, and computer-aided design and drafting (CADD) models that extend from simple wire-frame

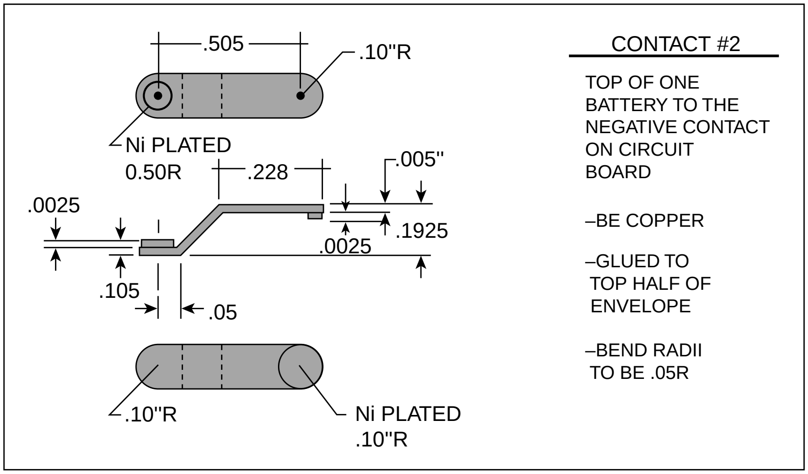



**Figure 9.1** Design information adjacent to a sketch of the designed object (after Ullman, Wood, and Craig (1990)). Note how clear and neat this sketch is, and that the notes are in easy-to-read block letters.

drawings (e.g., something very much like stick figures) through elaborate solid models (e.g., three-dimensional objects that include color and perspective).

In historical terms, drawing is the process of putting "marks on paper." These marks include both sketches and *marginalia*, that is, notes written in the margins. Sketches are of objects and their associated functions. Marginalia include notes in text form, lists, dimensions, and calculations. Thus, drawings enable a parallel display of information as they can be surrounded with adjacent notes, smaller pictures, formulas, and other pointers to ideas related to the object being drawn and designed. Putting notes next to a sketch is a powerful way to organize information, certainly more powerful than the linear, sequential arrangement imposed by the structure of sentences and paragraphs. Figure 9.1 illustrates some of these features in a sketch made by a designer working on the packaging for a battery-powered computer clock. (We will say more about the different kinds of sketches in Section 9.2.) The packaging consists of a plastic envelope and the electrical contacts. The designer has written down some manufacturing notes adjacent to the drawing of the spring contact. It would not be unusual for the designer to have scribbled modeling notes (e.g., "model the spring as a cantilever of stiffness . . . "), or calculations (e.g., calculating the spring stiffness from the cantilever beam model), or other information relating to the unfolding design. Thus, in such an annotated sketch, the designer can provide information about her assumptions, analyses that might be done, or even about how the device will be fabricated.

Marginalia of all sorts are familiar sights to anyone who has worked in an engineering environment. We often draw pictures and surround them with text and equations. We also draw sketches in the margins of documents, to elaborate a verbal description, to fortify understanding, to indicate more emphatically a coordinate system or sign convention. It is thus no surprise that sketches and drawings are essential to engineering design. In some fields (e.g., architecture), sketching, geometry, perspective, and visualization are acknowledged as the very underpinnings of the field.

In brief, graphic images are used to communicate with other designers, the client, and the manufacturing organization. Sketches and drawings:

- serve as a launching pad for a brand-new design;
- support the analysis of a design as it evolves;
- simulate the behavior or performance of a design;
- record the shape or geometry of a design;
- communicate design ideas among designers;
- ensure that a design is complete (as a drawing and its associated marginalia may remind us of still-undone parts of that design); and
- communicate the final design to the manufacturing specialists.

We will now present some more specifics about design sketches and the different kinds of engineering drawings.

## 9.2 SKETCHING

Sketching is a powerful tool in design because it enables us to convey our design ideas to others quickly and concisely. There are several types of sketches that designers routinely use to convey design information, including (see Figure 9.2):

- *Orthographic* sketches (Figure 9.2a) lay out the front, right and top views of a part.
- *Axonometric* sketches (Figure 9.2b) start with an axis, typically a vertical line with two lines 30° from the horizontal. This axis forms the corner of the part. The object is then blocked in using light lines, with the overall size first. Then vertical lines are darkened, followed by other lines. All lines in these sketches are either vertical or parallel to one of the two 30° lines. Details of the part are added last.
- *Oblique* sketches (Figure 9.2c) are probably the most common type of quick sketch. The front view is blocked in roughly first, depth lines are then added, and details such as rounded edges are added last.
- *Perspective* sketches (Figure 9.2d) are similar to oblique sketches in that the front view is blocked in first. Then a vanishing point is chosen and projection lines drawn from the points on the object to the vanishing point. The depth of the part is then blocked in using the projection lines. Finally, as in the other sketches, details are added to the part.

Each of these sketches provides a framework for putting design thoughts on paper. They may also serve as informational starting points for other design activities, for example, building physical models and prototypes. These quick sketches can also be enhanced if we pay attention to their proportions and if we annotate them:

- *Proportion control*: It is extremely useful to show the relative sizes of parts, components, or features in a sketch. That's why it is generally a good idea to sketch designs on graph paper, because it is easier to control those relative sizes using

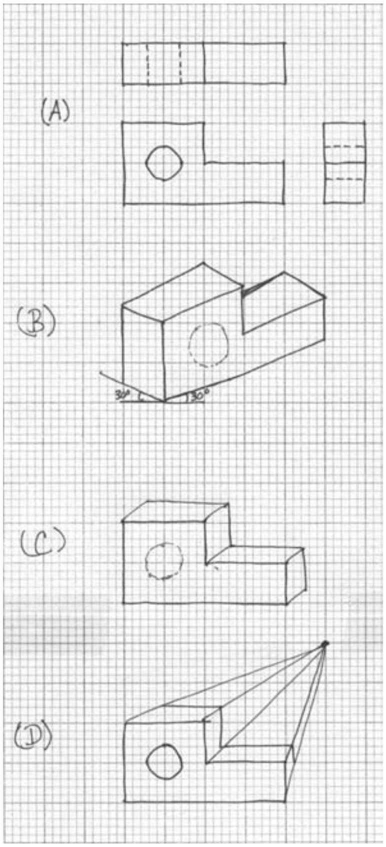



**Figure 9.2** Four types of sketches of the same object: (a) *orthographic*, (b) *axonometric*, (c) *oblique*, and (d) *perspective*.

the graph paper's grid, without having to make measurements with a ruler. It is also a good idea to think ahead of the components to be sketched before drawing actually begins: First "block in" the overall length and width of a part, to lay out the largest component first, and then add details.

- *Annotation*: Clear, easily read notes on sketches are tremendously useful in conveying the meaning behind the sketched ideas. We should make those annotations as clear and easy to read as possible. For example, it is generally useful to follow the familiar architectural style of using evenly spaced block letters; they are generally much clearer than cursive scribbles.

We saw a very good sketch in Figure 9.1: It was clear, drawn proportionately, and well annotated, using the block lettering we just suggested. Figure 9.3 shows two oblique sketches taken from team A's final report on the Danbury arm support project. These detailed design sketches have a rough visual appeal, and chances are that they took much more time to produce than sketches done early in the design process, where the focus is on getting out concepts quickly. We note, however, that the lettering in Figure 9.3 is uneven and hard to read in places. Block lettering would have taken no more time, and

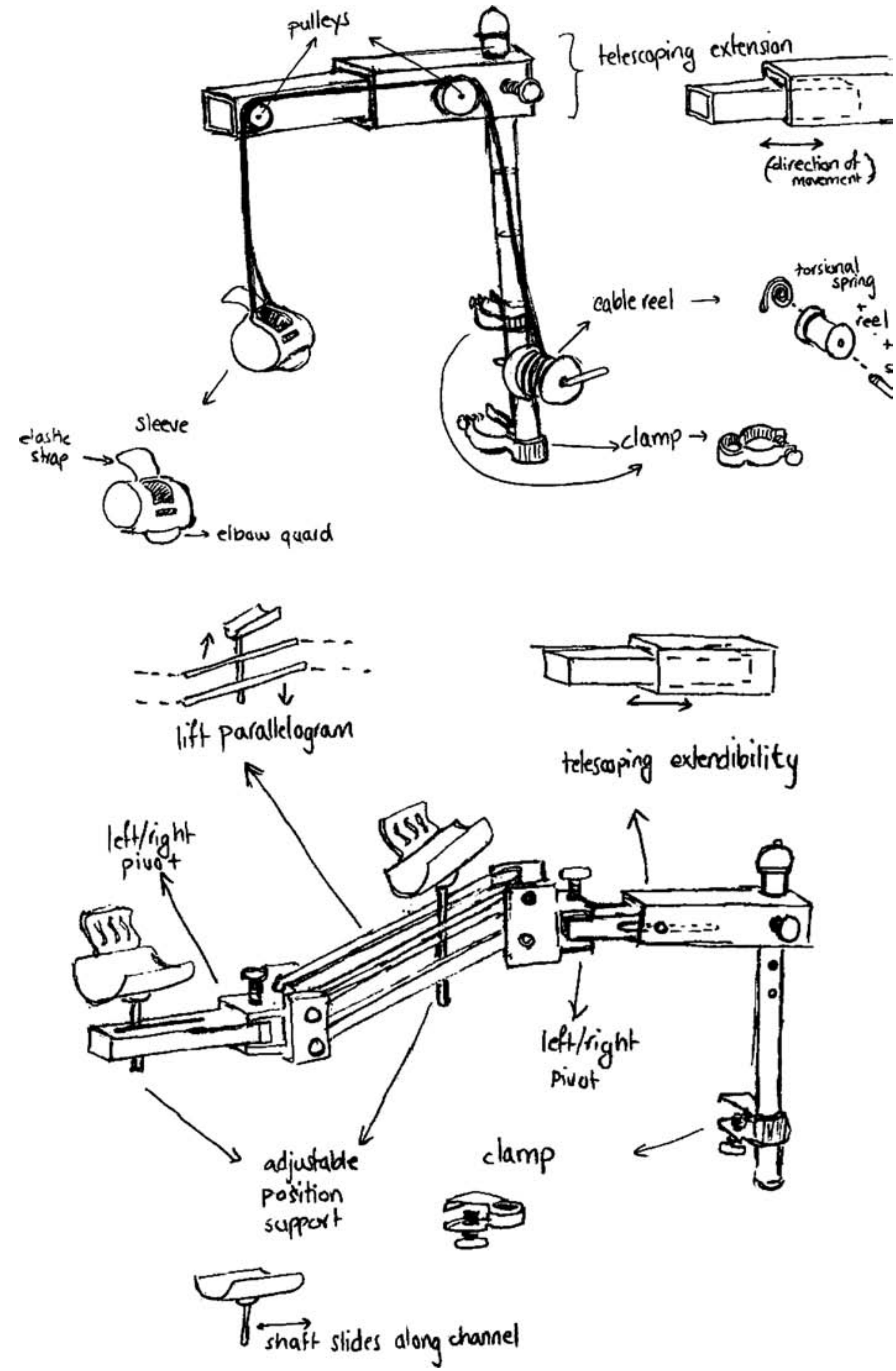



**Figure 9.3** Sketches of two of the design alternatives produced by team A for the Danbury arm support. Adapted from Attarian et al. (2007).

would have produced an overall visual image that is much easier to read and perhaps more impressive.

## 9.3 FABRICATION SPECIFICATIONS: THE SEVERAL FORMS OF ENGINEERING DRAWINGS

In addition to communicating with client(s) about a project, a design team must also communicate with the maker or manufacturer of the designed artifact. Often, the only "instructions" that the fabricator sees are those representations or descriptions of the designed object that are included in the final design drawings. This means that these representations must be complete, unambiguous, clear, and readily understood. The relevant question is: How do we ensure that the design as built will be exactly the design that we designed?

The answer is straightforward: When we communicate design results to a manufacturer, we must think very carefully about the *fabrication specifications* that we are creating in drawings, as well as those we write. This means we must ensure that our drawings are both appropriate to our design and prepared in accordance with relevant engineering practices and standards.

### 9.3.1 Design Drawings

There are several different kinds of drawings that are formally identified in the design process:

- *Layout drawings* are working drawings that show the major parts or components of a device and their relationship (see Figure 9.4). They are usually drawn to scale, do not show tolerances, and are subject to change as the design process evolves.

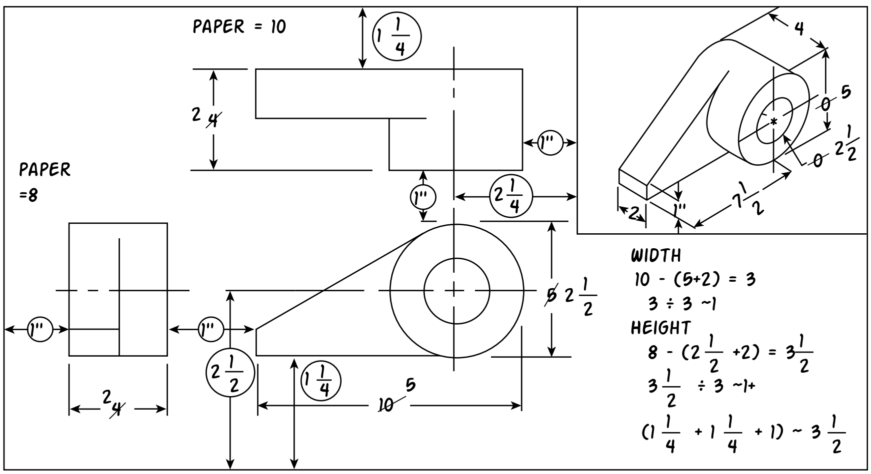



**Figure 9.4** A *layout drawing* that has been drawn to scale, does not show tolerances, and is certainly subject to change as the design process continues. Adapted from Boyer et al. (1991).

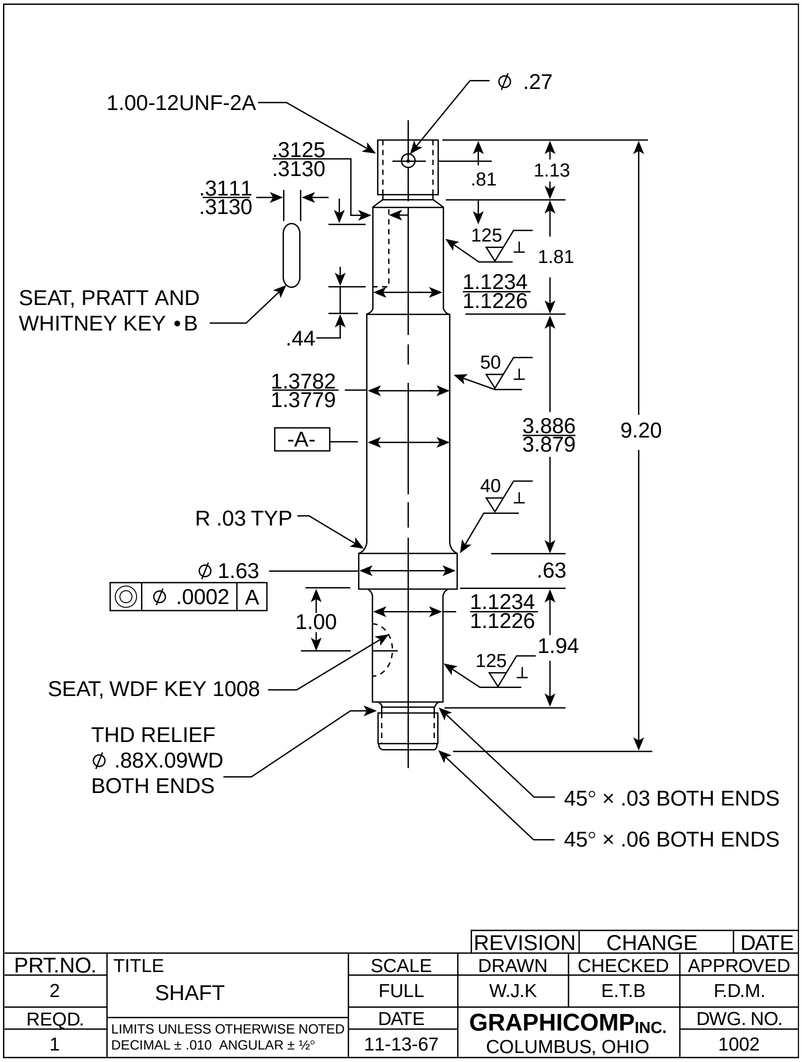



**Figure 9.5** A *detail drawing* that includes tolerances and that indicates materials and lists special processing requirements. It was drawn in conformance with ANSI drawing standards. Adapted from Boyer et al. (1991).

- *Detail drawings* show the individual parts or components of a device and their relationship (see Figure 9.5). These drawings must show tolerances, and they must also specify materials and any special processing requirements. Detail drawings are drawn in conformance with existing standards, and are changed only when a formal *change order* provides authorization.
- *Assembly drawings* show how the individual parts or components of a device fit together. An *exploded view* is commonly used to show such “fit” relationships

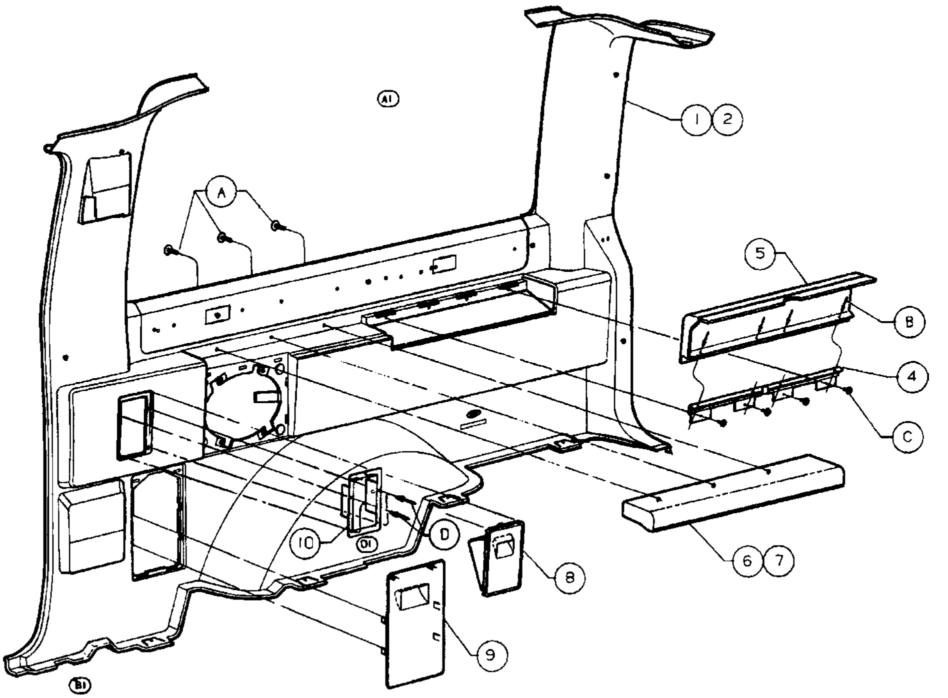



**Figure 9.6** This *assembly drawing* uses an *exploded view* to show how some of the individual parts of an automobile fit together. The individual components are identified by a part number or an entry on a bill of materials (that is not shown here). Adapted from Boyer et al. (1991).

(see Figure 9.6). We identify components by part numbers or entries on an attached *bill of materials*; they may include detail drawings if the major views in the detail drawings cannot show all of the required information.

In describing the three principal kinds of mechanical design drawings, we have used two technical terms that need definition: *tolerances* and *standards*. First, drawings show *tolerances* when they define the permissible ranges of variation in critical or sensitive dimensions. As a practical matter, it is impossible to make any two objects *exactly* the same. They may appear to be the same because of our limited ability to distinguish differences at extremely small or fine resolution. When we are producing many copies of the same thing intended to function in the same way, we must limit as best we can any variation from their ideally designed form. Tolerances formally prescribe these limits. The values chosen for tolerances on a specific part are often driven by the function of that part and by the manufacturing processes available.

Second, we have also noted the existence of drawing standards. *Standards* explicitly articulate the best current engineering practices in routine or common design situations. Thus, standards indicate performance bars that must be met for drawings (e.g., ASME Y14.5M–1994 *Dimensions and Tolerancing*), for the fire safety of buildings built within the United States (e.g., the *Life Safety Code* of the National Fire Protection Association), for boilers (e.g., the ASME *Pressure Vessel Code*), and so on. The American National Standards Institute (ANSI) serves as a clearinghouse for the individual standards written by professional societies (e.g., ASME, IEEE) and associations (e.g., NFPA, AISC) that govern various phases of design. ANSI also serves as the national spokesman for the United States in working with other countries and groups of countries (e.g., the European Union) to ensure compatibility and consistency wherever possible. A complete listing of U.S. product standards can be found in the *Product Standards Index*. The drawing standards specified in ASME Y14.5M–1994 are described in some depth in Appendix B.

### 9.3.2 Detail Drawings

We now turn to the requirements for detail drawings. These drawings are used to communicate the details of our design to the manufacturer or machinist. They must contain as much information as possible while being both as clear and as uncluttered as possible. Engineers and machinists have developed a system of standard symbols and conventions to meet this goal: *geometric dimensioning and tolerancing* (GD&T). These standards and conventions are described briefly here and in more detail in Appendix B.

Suppose we were designing a device, say a three-hole punch (see Exercise 9.1) or a simple desk lamp or a hammer, with the goal of having someone fabricate or make our design. What do we have to put down on paper, in both words and pictures, for us to be sure that the maker would fabricate exactly what we want? If we set down that description on paper and handed it to a friend or colleague, would she know exactly what we intend and want? This imaginary exercise is far more difficult than it sounds. In fact, imagine even a simpler version, akin to the way we introduced design problems. Supposed we had said to someone, “Please join this piece of metal to that piece of wood.” Is that a sufficient description of what we mean, for example, if we’re attaching steel rails to wooden railroad ties, or if we’re designing a clock to be housed in an elegant piece of grained maple?

The point is that as engineers we require common standards by which we can communicate our designs to those makers or machinists or fabricators who will actually make or build them. There are certain essential components that every drawing must have to ensure that it is interpreted as it is intended. These components include:

- standard drawing views;
- standard symbols to indicate particular items;
- clear lettering;
- clear, steady lines;
- appropriate notes, including specifications of materials;
- a title on the drawing;
- the designer’s initials and the date it was drawn;
- dimensions and units; and
- permissible variations, or tolerances.

Figure 9.7 displays a technical drawing that conforms to the GD&T system and ASME standards. It shows a screwdriver handle that Harvey Mudd students manufacture as

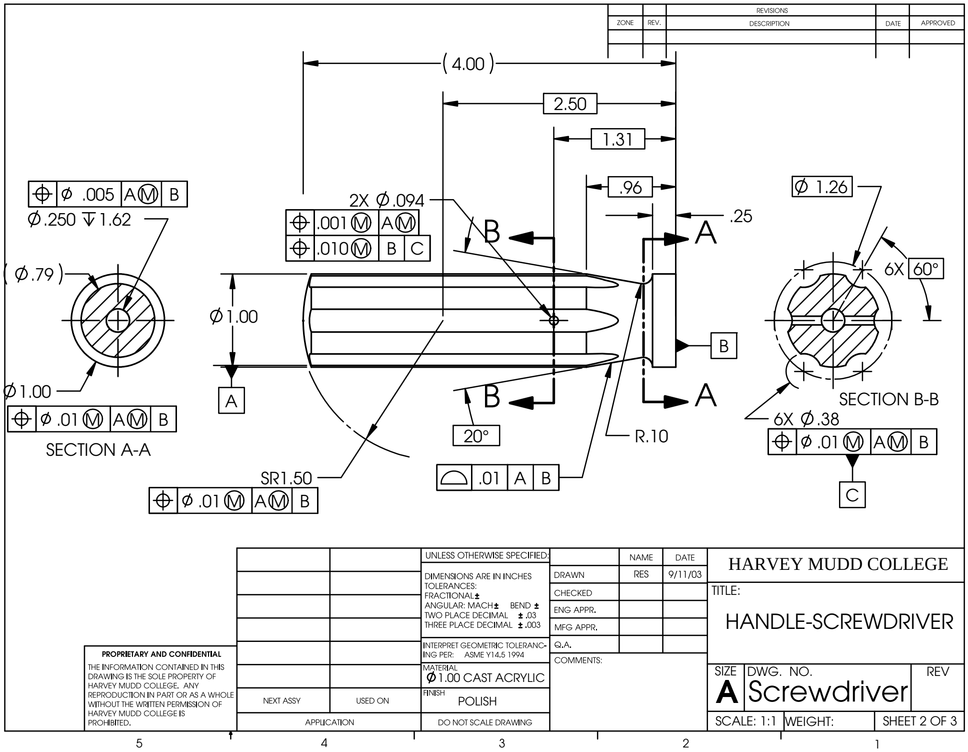
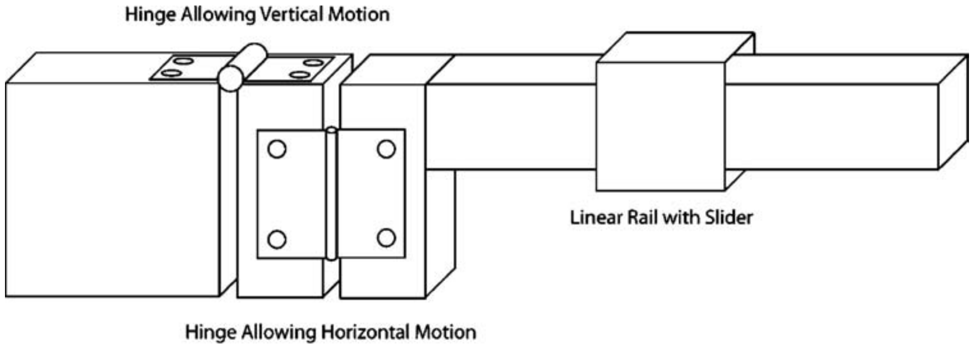



**Figure 9.7** A detail drawing of the handle of a screwdriver. This drawing uses a set of symbols and the particular placement of these symbols conveys information about the size and location of certain features of the screwdriver handle. In addition, the drawing contains information about the materials to be used, the finish of the part, the person who created it, and the date it was created. Drawing courtesy of R. Erik Spjut.

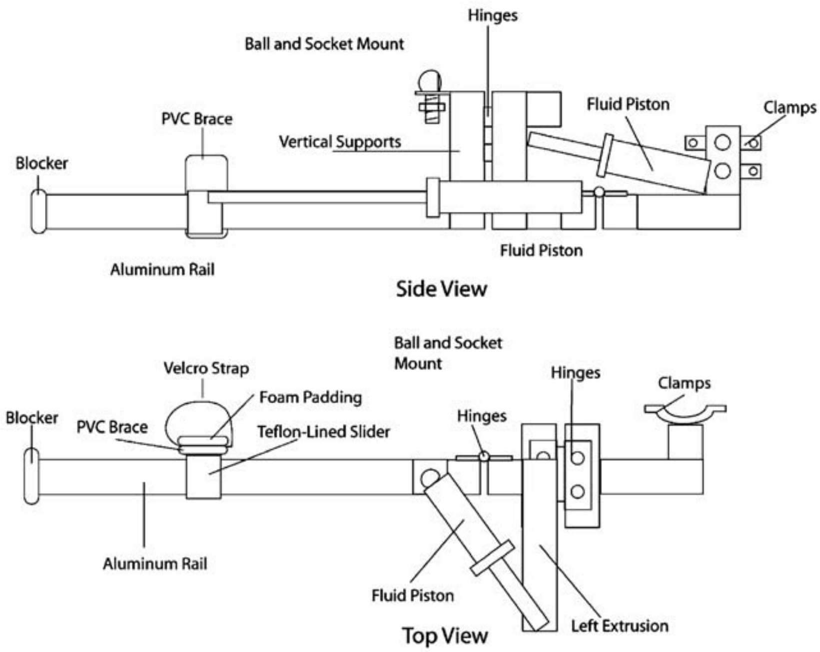
part of our introductory design course. Note the descriptive title, the date on the drawing, and the designer's initials. In addition, a note was included to specify the material as cast acrylic and that the finish should be polished. Dimensions and tolerances, specified in accord with GD&T standards and rules, are carefully detailed on the drawing. The various symbols that appear on the drawing are part of the GD&T language (and are described in more detail in Appendix B).

### 9.3.3 Some Danbury Arm Support Drawings

Design teams prepare drawings, often using drawing software, for technical reports and presentations. Figure 9.8 shows drawings created by team B for the Danbury arm support project. (We will show the corresponding prototypes in Chapter 10.) These drawings provide a lot of information and work well for presenting the work to clients. Note that team B used top and side views (think orthographic drawings!) to convey its design information. We should note, however, that these more formal sketches do not substitute for the detailed design drawings that must be produced in order to manufacture the design.




(a) Design team B's first design alternative




(b) Design team B's second design alternative

**Figure 9.8** Design drawings for the design alternatives produced by team B for the Danbury arm support. Adapted from Best et al. (2007).

## 9.4 FABRICATION SPECIFICATIONS: THE DEVIL IS IN THE DETAILS

The endpoint of a successful design project is the set of plans that form the basis on which the designed artifact will be built. It is not enough to say that this set of plans, which we have identified as the fabrication specifications, and which includes the final design drawings, must be clear, well organized, neat, and orderly. There are some very specific properties we want the fabrication specifications to have. They should be *unambiguous* (i.e., the role and place of each and every component and part must be unmistakable); *complete* (i.e., comprehensive and entire in their scope); and *transparent* (i.e., readily understood by the manufacturer or fabricator).

Fabrication specifications with these characteristics make it possible for the designed device or product to be built by someone totally unconnected to the designer or the design process. Remember that design must perform just as the designer intended because the designers may not be around to catch errors or to make suggestions, so the maker cannot turn around to seek clarification or ask on-the-spot questions. We have long passed the time when designers were also craftsmen who made what they designed. As a result, we can no longer allow designers much latitude or shorthand in specifying their design work because they are unlikely to be involved in the actual manufacture of the design result.

Fabrication specifications are normally proposed and written in the detailed design stage. Since our primary focus is conceptual design, we will not discuss fabrication specifications in depth. However, there are some aspects that are worth anticipating even early in the design process. One is that many of the components and parts that will be specified are likely to be purchased from vendors, for example, automobile springs, O-rings, DRAM chips, and so on. This means that a great deal of detailed, disciplinary knowledge comes into play. This detailed knowledge is often critically important to the lives of a design and its users. For example, many well-known catastrophic failures have resulted from inappropriate parts being specified, including the Hyatt Regency walkway connections, the Challenger O-rings, and the roof bracing of the Hartford Coliseum. The devil really is in the details!

Given that many parts and components can be bought from catalogs, while others are made anew, what sort of information must we, as designers, include in a fabrication specification? There are many kinds of requirements that we can specify in a fabrication specification, including:

- physical dimensions;
- materials to be used;
- unusual assembly conditions (e.g., bridge construction scaffolding);
- operating conditions (in the anticipated use environment);
- operating parameters (defining the artifact's response and behavior);
- maintenance and lifecycle requirements;
- reliability requirements;
- packaging requirements;
- shipping requirements;
- external markings, especially usage and warning labels; and
- unusual or special needs (e.g., must use synthetic motor oil).

This list of the different kinds of issues that must be addressed in a fabrication specification makes the point about our requirements for the properties of such a specification. The specification of the kind of spring action we see in a nail clipper might not seem a big deal, but the springs in the landing structure of a commercial aircraft had better be specified very carefully!

In the same way that there are different ways to write design requirements, we can anticipate different ways of writing fabrication specifications. When we specify a particular part and its number in a vendor's catalog, we are writing a *prescriptive* fabrication specification; when we specify a class of devices that do certain things, we are presenting a *procedural* fabrication specification; and when we leave it up to a supplier or the fabricator to insert something that achieves a certain function to a specified level, we are setting forth a *performance* fabrication specification.

## 9.5 FINAL NOTES ON DRAWINGS

There are many standards that define practices in the various engineering disciplines and domains. These are less likely to come into play in conceptual design, so we close our discussion of design drawings and fabrication specifications with a few general and philosophical observations.

First, different engineering disciplines use specific approaches that arose because of the ways in which these disciplines have grown and evolved, but they continue because of each discipline's needs. In mechanical design, for example, in order to make a complex piece that has a large number of components that fit together under extremely tight tolerances, we can't complete that design without constructing the sequence of drawings we described earlier. Therefore, we have to draw explicit depictions of the actual devices. In circuit design, on the other hand, both practice and technology have merged to the point that a circuit designer may be finished when she has drawn a circuit diagram, the analogy of a sketch of a spring–mass–damper. As designers, we must be aware that while there are habits and styles of thought that are common to the design enterprise, there are also practices and standards that are unique to each discipline—and it is our responsibility to learn and use them wisely.

We also want to reinforce the theme that some external, pictorial representation, in whatever medium, is absolutely essential for the successful completion of all but the most trivial of designs. These representations do not always take the form of a detailed engineering drawing. Some examples of pictorial representations that reflect the kinds of discipline-dependent differences we discussed above are: (1) the use of circuit diagrams to represent electronic devices, (2) the use of flowcharts to represent chemical-engineering-process plant designs, and (3) the use of block diagrams to represent control systems. These pictures, charts, and diagrams serve to extend our limited human abilities to flesh out and communicate the complicated pictures that exist solely within our minds.

Perhaps this is no more than a reflection of a more accurate translation of a favorite Chinese proverb, "One showing is worth a hundred sayings." It may also reflect the German proverb, "The eyes believe themselves; the ears believe other people." A good sketch or rendering can be very persuasive, especially when a design concept is new or controversial. A drawing is an excellent way to group information because its nature allows us to put additional information about an object in an area adjacent to its "home" in a drawing. We can do this for the design of a complex object as a whole, or on a more localized, part-by-part basis.

## 9.6 NOTES

*Section 9.1*: Much of the discussion of drawing is drawn from Ullman, Wood, and Craig (1990) and Dym (1994). The listing of kinds of design drawings is adapted from Ullman (1997). The drawings shown in Figures 9.4–9.6 are adapted from Boyer et al. (1991).

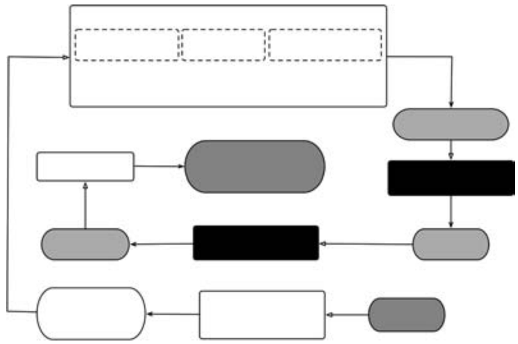
*Section 9.2*: Our brief overview of the basics of dimensioning and tolerancing here and in Appendix B has relied on the information found in ASME (1994), TDCA (1996), Goetsch (2000), and Wilson (2005).

*Section 9.5*: The Chinese and German proverbs are from Woodson (1966).

CHAPTER 10

# PROTOTYPING AND PROOFING THE DESIGN

*Here's my design; how well does it work?*



DESIGN RESULTS can be communicated in several ways. In this chapter we focus on how to translate our design ideas into models and prototypes that can be used to test our design concepts and communicate our ideas to the client. Often the first step in such a process involves sketching or drawing our design, as we discussed in Chapter 9, because we can use these representations to create the prototype or model. One useful tool for this process is a three-dimensional (3D) representation of the designed object in a software program such as Cleo Elements/Pro$^{TM}$ (formerly ProEngineer) or SolidWorks$^{TM}$. This 3D representation can then be used: (1) as an input to a computational modeling program to simulate the design's performance under specified conditions; (2) as an input into a variety of rapid prototyping technologies, such as 3D printing; (3) to generate detailed engineering drawings of the design; and (4) to guide the tool path in computer numerical-controlled (CNC) machining. We now describe a few common modeling and prototyping tools.

## 10.1 PROTOTYPES, MODELS, AND PROOFS OF CONCEPT

For many engineering projects, we want to build three-dimensional, physical realizations of our concepts and designed artifacts. There are several versions of physical things that could be made—including prototypes, models, and proofs of concept—and they are often made by their designers.

*Prototypes* are "original models on which something is patterned." They are also defined as the "first full-scale and usually functional forms of a new type or design of a construction (such as an airplane)." In this context, prototypes are working models of designed artifacts. They are tested in the same operating environments in which they're expected to function as final products. It is interesting that aircraft companies routinely build prototypes, while rarely, if ever, does anyone build a prototype of a building.

A *model* is "a miniature representation of something," or a "pattern of something to be made," or "an example for imitation or emulation." We use models to *represent* some devices or processes. They may be paper models or computer models or physical models. We use them to illustrate certain behaviors or phenomena as we try to verify the validity of an underlying (predictive) theory. Models are usually smaller and made of different materials than are the original artifacts they represent, and they are typically tested in a laboratory or in some other controlled environment to validate their expected behavior.

A *proof of concept*, in this context, refers to a model of some part of a design that is used specifically to test whether a particular concept will actually work as proposed. Doing proof-of-concept tests means doing controlled experiments to prove or disprove a concept.

### 10.1.1 Prototypes and Models Are Not the Same Thing

The definitions of prototypes and models sound enough alike that it prompts a question: Are prototypes and models the same thing? The answer is, "Not exactly." The distinctions between prototypes and models may have more to do with the intent behind their making and the environments in which they are tested than with any clear dictionary-type differences. Prototypes are intended to demonstrate that a product will function as designed, so they are tested in their actual operating environments or in similar, uncontrolled environments that are as close to their relevant "real worlds" as possible. Models are intentionally tested in controlled environments that allow the model builder (and the designer, if they are not the same person) to understand the particular behavior or phenomenon that is being modeled. An airplane prototype is made of the same materials and has the same size, shape, and configuration as those intended to fly in that series (i.e., Boeing 747s or Airbus 310s). A model airplane would likely be much smaller. It might be "flown" in a wind tunnel or for sheer enjoyment, but it is not a prototype. A prototype is the first of its kind; a model represents a device or a process.

Engineers often build models of buildings, for example, to do wind-tunnel testing of proposed skyscrapers, but these models are not prototypes. Rather, building models used in a wind-tunnel simulation of a cityscape with a new high-rise are essentially toy building blocks that are meant to imitate the skyline. They are not buildings that work in the sense of aircraft prototypes that actually fly. So, why do aeronautical engineers build prototype airplanes, while civil engineers do not build prototype buildings? What do they do in other fields?

### 10.1.2 Testing Prototypes and Models, and Proving Concepts

We introduced testing in discussing both models and prototypes. In design the type of testing that is often most important is *proof of concept* testing in which a new concept, or a particular device or configuration, can be shown to work in the manner in which it was designed. When Alexander Graham Bell successfully summoned his assistant from another room with his new-fangled gadget, Bell had proven the concept of the telephone. Similarly, when John Bardeen, Walter Houser Brattain, and William Bradford Shockley successfully controlled the flow of electrons through crystals, they proved the concept of the solid-state electronic valve, know as the transistor, that replaced vacuum tubes. Laboratory demonstrations of wing structures and building connections can also be considered as proof-of-concept tests when they are used to validate a new wing structure configuration or a new kind of connection. In fact, even market surveys of new products—where samples are mailed out or stuffed into sacks in the Sunday papers—can be conceived of as proof-of-concept tests that test the receptivity of a target market to a new product.

Proof-of-concept tests are scientific endeavors. We set out reasoned and supported hypotheses that are tested and then validated or disproved. Turning on a new artifact and seeing whether or not it "works" is not a proper proof-of-concept demonstration. An experiment must be designed, with hypotheses to be disproved if certain outcomes result. Remember that prototypes and models differ in their underlying "reasons for being" and in their testing environments. While models are tested in controlled or laboratory environments, and prototypes are tested in uncontrolled or "real-world" environments, both are *controlled tests*. Similarly, when we are doing proof-of-concept tests, we are doing controlled experiments in which the failure to disprove a concept may be key.

For example, suppose we had chosen Mylar containers as our new juice product and designed them to withstand shipping and handling, both in the manufacturing plant and in the store. If we think of all of the things that could go awry (e.g., stacks of shipping pallets that could topple) and analyze the mechanics of what happens in such incidents, then we might conclude that the principal design criterion is that the Mylar containers should withstand a force of X Newtons (N). We would then set up an experiment in which we apply a force of X N, perhaps by dropping the containers from a properly calculated height. If the bags survived that drop, we could say that they'd likely survive shipping and handling. However, we could not absolutely guarantee survival because there is no way we could completely anticipate every conceivable thing that might happen to a beverage-filled Mylar container. On the other hand, if the Mylar container fails a properly designed drop test, we can then be certain that it will not survive shipping and handling, and so our concept is disproved. The National Aeronautics and Space Administration (NASA) conducted a similar proof-of-concept test for gas-filled shock absorbers for one of the Mars landers. There are potential issues of legal liability involved in product testing (for example, for how much nonstandard use of a product can a manufacturer be held responsible?), but they are beyond the scope of this text.

Prototypes, models, and proof-of-concept testing have different roles in engineering design because of their intents and test environments. These distinctions must be borne in mind while planning them for the design process.

### 10.1.3 When Do We Build a Prototype?

The answer is, "It depends." The decision to build a prototype depends on a number of things, including: the size and type of the design space, the costs of building a prototype, the ease of building that prototype, the role that a full-size prototype might play in ensuring the widespread acceptance of a new design, and the number of copies of the final artifact that are expected to be made or built. Aircraft and buildings provide interesting illustrations because of ample commonalities and sharp differences. The design spaces of both aircraft and high-rises are large and complex. There are millions of parts in each of these examples, and so many design choices are made along the way. The costs of building both airplanes and tall buildings are also quite high. In addition, at this point in time, we have ample experience with both aeronautical and structural technologies, so that we generally have a pretty good idea of what we're about in these two domains. So, again, why prototype aircraft and not prototype buildings? In fact, don't the complexity and expense of building even a prototype aircraft argue directly against the idea of building such prototypes?

Notwithstanding all of our past experience with successful aircraft, we build prototypes of airplanes because, in large part, the chances of a catastrophic failure of a "paper design" are still unacceptably high, especially for the highly regulated and very competitive commercial airline industry that is the customer for new civilian aircraft. That is, we are simply not willing to pay the price of having a brand new airplane take off for the first time with a full load of passengers, only to watch hundreds of lives being lost—as well as the concomitant loss of investment and of confidence in future variants of that particular plane. It is in part an ethical issue, because we do bear responsibilities for technical decisions when they impinge upon our fellow humans. It is also in part an economic issue, because the cost of a prototype is economically justifiable when weighed against potential losses. When we build prototypes of airplanes those particular planes are not simply thrown away as "losses" after testing; they are retained and used as the first in the series of the many full-size designs that are the rest of the fleet of that kind of airplane.

Buildings do fail catastrophically, during and after construction. However, this occurs so rarely that there is little perceived value in requiring prototype testing of buildings before occupancy. Building failures are rare in part because high-rises can be tested, inspected, and experienced gradually, as they are being built, floor by floor. The continuous inspection that takes place during the construction of a building, from the foundation on up, has its counterpart in the numerous inspections and certifications that accompany the manufacture and assembly of a commercial airliner. But the maiden flight of an airplane is a binary issue, that is, the plane either flies or it doesn't, and a failure is not likely to be a graceful degradation!

Another interesting aspect of comparing the design and testing of airplanes to that of buildings has to do with the number of copies being made. We have already noted that prototype aircraft are not discarded after their initial test flights; they are flown and used.

In fact, airframe manufacturers are in business to build and sell as many copies of their prototype aircraft as they can, so engineering economics plays a role in the decision to build a prototype. The economics are complicated because the manufacturing cost of the first plane in a series is very high. Technical decisions are made about the kinds of tooling and the numbers of machines needed to make an airplane, and economic trade-offs between the anticipated revenue from the sale of the airplane and the cost the manufacturing process are evaluated. We will address some manufacturing cost issues in Chapter 13.

Another lesson we can learn from thinking about buildings and airplanes is that there is no obvious correlation between the size and cost of prototyping—or the decision to build a prototype—and the size and type of the design space. And while it might seem that the decision to build a prototype might be strongly influenced by the relative ease of building it, the aircraft case shows that there are times when even costly, complicated prototypes must be built. On the other hand, if it is cheap and easy to do, then it generally would seem a good idea to build a prototype. There certainly are instances where prototypes are commonplace, for example, in the software business. Long before a new program is shrink-wrapped and shipped, it is alpha- and beta-tested as early versions are prototyped, tested, evaluated, and, hopefully, fixed.

If there is a single lesson about prototypes, it is that the project schedule and budget should reflect plans for building them. More often than not a prototype is required, although there may be instances in which resources or time are not available. In weapons development contracts, for example, the U.S. Department of Defense virtually always requires that design concepts be demonstrated so that their performance can be evaluated before costly procurements. At the same time, it is interesting that aircraft companies (and others) are demonstrating that advances in computer-aided design and analysis allow them to replace some elements of prototype development with sophisticated simulation.

Sometimes we build prototypes of parts of large, complex systems to use as models to check how well those parts behave or function. For example, structural engineers build full-size connections, say, at a point where several columns and beams intersect in a geometrically complicated way, and test them in the laboratory. Similarly, aeronautical engineers build full-size airplane wings and load them with sandbags to validate analytical models of how these wing structures behave when loaded. A prototype of a part of a larger artifact is built in both instances, and then used to model behavior that needed to be understood as part of completing the overall design. We use prototypes to demonstrate functionality in the real world of the object being designed, and we use models in the laboratory to investigate and validate behavior of a miniature or of part of a large system.

## 10.2 BUILDING MODELS AND PROTOTYPES

There are principles behind and guidelines for building and testing prototypes and models. The important questions we must ask are: What do we want to learn from the model or prototype? Who is going to make it? What parts or components can be bought? How, and from what, is it going to be made? How much will it cost? We have already answered the first question in Section 10.1, but we should keep our answers in mind as we turn to the implementation details of (actually) building a model or a prototype.

### 10.2.1 Who Is Going to Make It?

We have two basic choices when we want a model or prototype: we make and/or assemble it in-house or we outsource it. With sufficient time and money, we can make just about anything we want, from an application-specific integrated circuit to a computational-fluid-dynamics (CFD) model of a pressure relief valve to a pilot-scale oil refinery. Three major factors enter into the decision about who makes our model: expertise, expense, and time.

Many companies and schools keep machinists, electronics technicians, and programmers on staff for building models. They often have facilities that individuals can use to build prototypes, and some schools even require students to learn how to use such facilities. However, it is rather rare to have facilities or expertise for making intricate or tight-tolerance items. Thus, we first want to ask if anyone on our design team has the necessary expertise, or is willing to learn. We should identify expertise and facilities that are available in-house. If the needed expertise is missing in-house, we should plan on having parts or components outsourced.

Time and cost are usually intertwined. If we need a part "yesterday," it's not likely that we can outsource it without a spending a lot of money, but we may be able to go to our own machine shop and machine it in an hour. However, engineers aren't always allowed to use their company's machine shop, so we might have to convince a machinist to do it. Then the likelihood of getting it done right away will depend on the machinist's workload and her willingness to oblige. We can see that it is a good idea to cultivate good relationships with a company's (or a school's!) machinists and technicians. It is always a good idea to try to give them meaningful lead times, to ask their advice often, and to not ask for things that are silly or impossible. Treat technicians and machinists—and indeed, all staff—as professionals and as equals. In addition to being the right thing to do, this will make it much more likely that we can get their help when we need it.

Sometimes, especially for specialized items, it may still be cheaper and/or faster to have specific items outsourced. For example, gears and printed circuit boards are likely to be gotten cheaper and faster when outsourced.

If we do something outsourced, it will go much more smoothly if we prepare detailed specifications for what we are outsourcing. That might include properly drawn, toleranced, and specified mechanical drawings for machined or manufactured parts (see Chapter 9); accurate and double-checked Gerber files for printed circuit boards; or complete and correct part numbers for parts or components.

### 10.2.2 Can We Buy Parts or Components?

There are a lot of parts and components that are best bought from suppliers, unless we happen to be in the business of designing and making those particular items. For example, it is rarely worth the time, equipment, and expense to make screws or transistors: Common mass-produced items should always be bought—although it also a good idea to check your institution's stockroom(s) to see whether the parts are already available. Fasteners, such as nails, nuts, bolts, screws, and retaining rings should almost always be bought, as should common mechanical parts or devices such as pulleys, wheels, gears, casters, transmissions, and hinges. Similarly, electronic, electromechanical, and optical components such as resistors, capacitors, integrated circuits, electric motors, solenoids, light-emitting diodes (LEDs), lenses, and photodiodes can be bought.

**TABLE 10.1 A short list of suppliers and their URLs for a variety of products that might be useful in building models and prototypes**

| What Is Being Sought? | Supplier | URL |
|---|---|---|
| Materials, mechanical items | McMaster-Carr | http://www.mcmaster.com/ |
| | Grainger | http://www.grainger.com/ |
| Electronic supplies | Digi-Key | http://www.digikey.com/ |
| | Newark | http://www.newark.com/ |
| | Mouser | http://www.mouser.com/ |
| Optical, opto-mechanical, electro-optical components | Thorlabs | http://www.thorlabs.com/ |
| | Newport | http://www.newport.com/ |
| Suppliers (i.e., catalogs of other suppliers) | Thomas Register | http://www.thomasnet.com/ |
| | Global Spec | http://www.globalspec.com/ |

The Internet and the widespread use of *just-in-time* manufacturing have dramatically changed the ease with which we can find suppliers and parts. We show a list of some widely available suppliers and their URLs in Table 10.1. Many companies allow online ordering and provide fast (i.e., overnight) shipping of in-stock items. Some companies will ship both small lots and large quantities, and some websites provide excellent search capabilities (assuming we know the proper term for the part we want) and display real-time inventories.

It really does pay to spend some time searching for parts and components. An experienced engineer or a librarian can often be very helpful as we search. We should note the prices or cost when we find what we want because that will come in handy for our budget. We can likely find many, or even all, of our prototype components already available online or in a store—and it's far better to learn that before we start building, instead of just as we finish.

### 10.2.3 How, and from What, Will the Model/Prototype Be Made?

Now we turn to basic principles and best practices for making, machining, and assembling mechanical parts. The carpenter's maxim is: *measure twice, cut once*. We should create detailed, annotated drawings and plans *before* we start cutting or machining: it will pay huge dividends in ensuring that things fit together the very first time they are assembled, and will surely minimize reworking the problem or the parts.

Detailed plans should include a *bill of materials*, which we will describe in Chapter 13 as an important aspect of engineering economics in design. It is easier to construct a bill of materials during a virtual or paper assembly, and so identify all of the necessary parts *before* everything has been done. Then we can check on the availability of the parts to help our scheduling. It may also be useful to construct a *process router*: a list of instructions for how the prototype is to be fabricated and assembled.

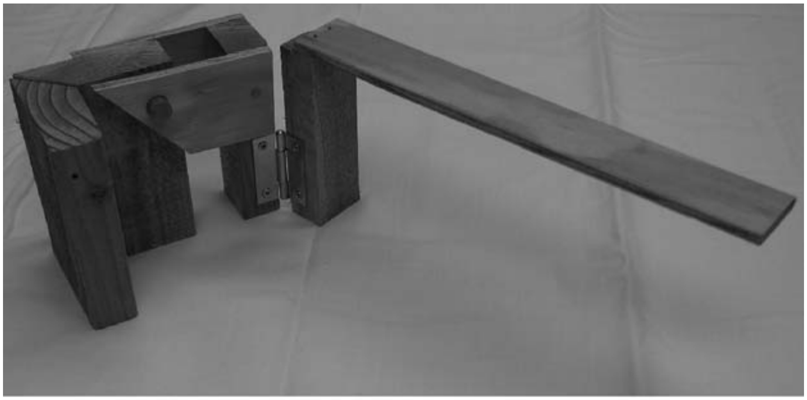


Photo by Clive L. Dym

**Figure 10.1** Team B's wooden model of their design for the Danbury arm support. The model was constructed to demonstrate and clarify the dual articulation of the arm motions that were to be supported.

True prototypes are typically made from the same materials that are intended for the final design. Of course, those materials may change for the final design as a response to what was learned from the prototype. A model, on the other hand, can be constructed from whatever material will assist in answering the questions for which the model was designed to pose. The most common materials for model construction are paper, cardboard, wood, plywood, polymers (such as PVC, ABS, polystyrene, and acrylic), aluminum, and mild steel. (See Appendix A for more details on materials choices.)

Figure 10.1 shows the model that team B built, and Figure 10.2 shows prototypes from both Danbury project teams. We note that the model was built of wood because one team member, being skilled in woodworking, realized that he could clearly demonstrate team B's arm concept with that model. The two prototypes were built largely of aluminum because both teams wanted to leave their prototypes behind so that Jessica and other Danbury students could actually—and safely—use them.

There are many options for constructing prototypes and models. We can construct mockups of a design from 2D shapes, machine parts directly, CNC machine parts, or use rapid prototyping technologies. Our choice of which of these to use depends on the cost, timing, and complexity of our design:

- *Mock-ups*: One option for making basic models or prototypes is to construct a mock-up of a 3D part from 2D cutouts. These 2D parts can be made using a vinyl cutter or a laser cutter, and parts are then assembled into 3D mock-ups of a design. Materials used for these mock-ups might be foam, thin plastic, or wood.
- *Machining*: We may have the option of machining parts or all of our prototypes ourselves in a machine shop. Typically, there are separate machine shops for woodworking and metalworking. Woodworking machines include *drill presses* for making holes, *band saws* for cutting at various angles, and *lathes* for reducing diameter and creating parts with symmetric curved surfaces. A metal shop includes

(a)

Photo by Clive L. Dym

(b)

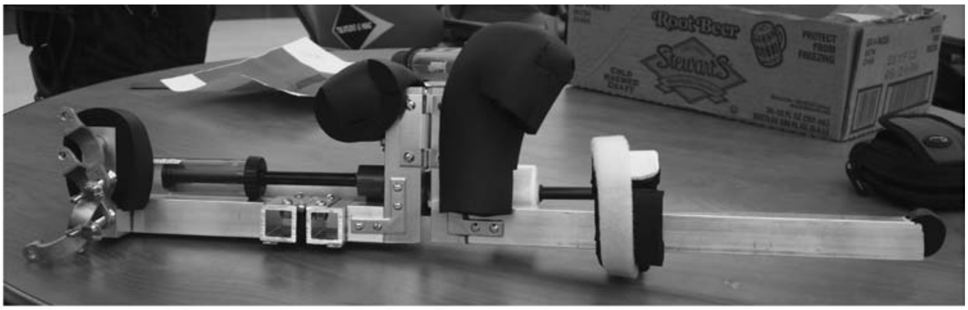



Photo by Clive L. Dym

**Figure 10.2** Prototypes of the Danbury arm support, made by teams A (a) and B (b).

*lathes* for reducing diameter of a part, tapping a hole or facing an end, and *mills* for creating slots, holes, and flat surfaces. Basic techniques for shaping and joining machined materials are described in Appendix A.

If our part is particularly complex, it may be simpler to use CNC machining to produce it. CNC machines range in size and cost and can produce very small objects or very large objects in a range of materials. These machines use the 3D computer-aided design (CAD) models and/or drawings of your part to create a step-by-step machine log in a software program such as Mastercam™, which is then input into the CNC machine. These machines can achieve much higher tolerances and can produce much more intricate parts than machining by hand.

If we have the option and choose to machine the prototype ourselves, we need to remember that *safety in the workshop is critically important*. Power tools can easily cause dismemberment and death. A moment's inattention can lead to a permanent change in lifestyle and career. For more details on common machine shop safety practices, see Appendix A.

- *Rapid prototyping technologies*: Rapid prototyping technologies have emerged in recent years as relatively fast and cheap ways to fabricate prototypes that would otherwise need to be injection molded. Rapid prototyping techniques use 3D CAD models as inputs, and convert these 3D files into thin 2D layers to build the 3D part. Rapid prototyping technologies include stereo-lithography and selective laser sintering, which involve using a laser to harden either a resin bath or a polymer powder in a particular configuration to build each layer.

Another technique is fused deposition modeling, in which a heated filament of a particular material is squeezed out of a tube one layer at a time onto a stage. The stage is then moved down a fixed increment and another layer completed. Fused deposition modeling utilizes standard engineering thermoplastics, most commonly acrylonitrile–butadiene–styrene (ABS), but more than one material can be used at once. Thermoplastics are used to make light, rigid, molded products. ABS is impact resistant and tough, making the parts produced from fused deposition modeling structurally functional. Parts made by this method can be sanded and machined postprinting to finish the part. Most commercial 3D printers using this technology have build resolutions of 0.254 mm (0.010 in), but the latest technology available has a build resolution of 16 microns (0.0006 in) and the capability of using multiple materials to build parts, with material properties ranging from rubber to thermoplastics.

3D printed parts can be very useful for communicating ideas to our client and can be created relatively quickly, but there are some drawbacks. Figure 10.3 shows both a 3D printed version of a screwdriver and its hand-machined counterpart. Both parts were made by HMC students in the design realization component of the E4 design course. The 3D printed version took 4 h to complete, while the machined part can take 15–20 h for a new machinist. The 3D printed version can be very useful for communicating overall size of the

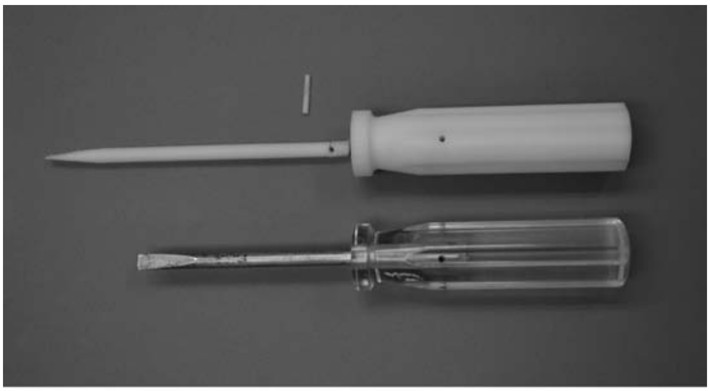


Photo by Elizabeth J. Orwin

**Figure 10.3** 3D printed (top) and machined (bottom) versions of the HMC E4 screwdriver. Note that without sanding, the 3D printed blade cannot fit into the handle because of print resolution limits.

design and ergonomic features, such as the depth of the grooves in the handle. However, the printed blade and handle do not fit together, as the print resolution did not meet the required tolerances. In addition, small pieces, such as the pin, are very brittle and can be easily broken. It is easier to achieve specified tolerances by machining a part by hand, and if our tolerances are really tight, we should consider CNC machining.

### 10.2.4 How Much Will It Cost?

We will describe some aspects of estimating cost in Chapter 13, but it is worth noting here that it is wise to plan for miscalculations of prices and of the numbers of parts and amounts of materials that we will use. Prices always seem to go up between the time items are priced and the time they are bought. And we often forget to include sales tax and shipping in our costs. So it's good to leave a margin for error, say 10–15%, especially if this is a first-ever model-building project. We should also be careful to ensure that any big-budget items really meet our needs *before* we order them: A 10% reserve won't help if the $100 item in the $125 budget is the wrong one.

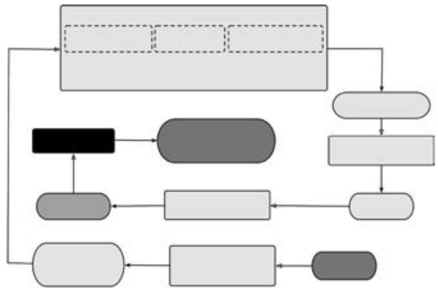
## 10.3 NOTES

*Section 10.2*: The definitions quoted at the beginning of this discussion are from *Webster's Tenth Collegiate Dictionary* (Mish 1993).

# CHAPTER 11

# COMMUNICATING DESIGNS ORALLY AND IN WRITING

*How do we let our client know about our solutions?*



REPORTING IS an essential part of a design project: We have not completed our project if we have not communicated our work and findings to our client and to other stakeholders the client may designate. We communicate final design results in several ways, including oral presentations, final reports (that may include design drawings and/or fabrication specifications), and prototypes and models. In this chapter we first consider some common guidelines for all reporting modes, and then we look at oral presentations and at final technical reports.

The primary purpose of such communication is to inform our client *about the design*, including explanations of how and why this design was chosen over competing design alternatives. It is most important that we convey the *results* of the design process. The client is probably not interested in the history of the project or in the design team's internal workings, and so we should ensure that final reports and presentations are *not* narratives or chronologies of our work. Rather, our presentations and reports should be lucid descriptions of design *outcomes*, as well as the processes with which those outcomes were achieved.

## 11.1 GENERAL GUIDELINES FOR TECHNICAL COMMUNICATION

There are some basic elements of effective communication that apply to writing reports, giving oral presentations, and even providing informal updates to your client. Thomas Pearsall summarized these common concepts as the seven principles of technical writing but they clearly apply more generally:

1. *Know your purpose.*
2. *Know your audience.*
3. *Choose and organize the content around your purpose and your audience.*
4. *Write precisely and clearly.*
5. *Design your pages well.*
6. *Think visually.*
7. *Write ethically!*

While Pearsall devoted more than one-half of his book to these principles, we will summarize them here as a prelude to the rest of this chapter.

***Know your purpose***. This is the writing analog of understanding objectives and functions for a designed artifact. Just as we want to understand what the designed object must be and must do, we need to understand the goals of a report or presentation. In many cases design documentation seeks to inform the client about the features of a selected design. In other cases the design team may be trying to persuade a client that a design is the best alternative. In still other cases a designer may wish to report how a design operates to users, whether beginners or highly experienced ones. If you don't know what purpose you are trying to serve with your writing or presentation, you may not produce anything or serve any purpose.

***Know your audience***. We have all sat through lectures where we didn't know what was going on or where the material was so simple that we already knew it. We can often take some action once we realize that the material is not set at an appropriate level. Similarly, when documenting a design, it is essential that a design team structure its materials to its targeted audience. Thus, the team should ask questions like, "What is the technical level of the target audience?" and "What is their interest in the design being presented?" Taking time to understand the target audience will help ensure that its members appreciate your documentation. Sometimes you may prepare multiple documents and briefings on the same project for different audiences. For example, it is quite common for designers to close projects out with both a technical briefing and a management briefing. It is also common for designers to confine calculations or concepts that are of limited interest to a report's primary audience to specific sections of their reports, usually appendices.

***Choose and organize the content around your purpose and your audience***. Once we are sure of the purpose of the report or presentation and its target audience, it only makes sense to try to select and organize its content so that it will reach its intended target. The key element is to structure the presentation to best reach the audience. In some cases, for example, it is useful to present the entire process by which the design team selected an alternative. Other audiences may only be interested in the outcome.

There are many different ways to organize information, including going from general concepts to specific details (analogous to deduction in logic), going from specific details to general concepts (analogous to induction or inference), and describing devices or systems.

Once an organizational pattern is chosen, no matter which form is used, the design team should translate it into a written outline. This allows the team to develop a unified, coherent document or presentation while avoiding needless repetition.

***Write precisely and clearly***. This particular guideline sounds like "use common sense," that is, do something that everyone wants to do but few achieve. There are, however, some specific elements that seem to occur in all good writing and presentations. These include effective use of short paragraphs that have a single common thesis or topic; short, direct sentences that contain a subject and a verb; and active voice and action verbs that allow a reader to understand directly what is being said or done. Opinions or viewpoints should be clearly identified as such. These elements of style should be learned so that they can be correctly applied. Young designers may have practiced these skills more in humanities and social science classes than in technical courses. This is acceptable, and even welcome, so long as the designer remembers that the goals of both technical and nontechnical communications remain the same.

***Design your pages well***. Whether writing a technical report or organizing supporting materials for a verbal briefing or presentation, effective designers utilize the characteristics of their media wisely. In technical reports, for example, writers judiciously use headings and subheadings, often identified by different fonts and underlining, to support the organizational structure of the report. A long section divided into several subsections helps readers to understand where the long section is going, and it sustains their interest over the journey. Selecting fonts to highlight key elements or to indicate different types of information (such as new, important terms) guides the reader's eye to key elements on the page. Tables should be treated as a single figure and should not be split over a page break. White space on a page helps keep readers alert and avoids a forbidding look in documents.

Similarly, careful planning of presentation support materials such as slides and transparencies can enhance and reinforce important concepts or elements of design choices. Using fonts that are large enough for the entire audience to see is an obvious, but often overlooked, aspect of presentations. Just as white space on a page invites readers to focus on the text without being distracted, simple and direct slides encourage readers to listen to the speaker without being distracted visually. Thus, text on a slide should present succinct concepts that the presenter can amplify and describe in more detail. A slide does not have to show every relevant thought. It is a mistake to fill slides with so many words (or other content) that audiences have to choose between reading the slide and listening to the speaker, because then the presenter's message will almost certainly be diluted or lost.

***Think visually***. By their very nature, design projects invite visual thinking. Designs often start as sketches, analyses often begin with free-body or circuit diagrams, and plans for realizing a design involve graphics such as objectives trees and work breakdown structures. Just as designers often find that visual approaches are helpful to them, audiences are helped by judicious use of visual representation of information. These can range from the design tools discussed throughout this book, to detailed drawings or assembly drawings, to flow charts and cartoons. Even tables present an opportunity for a design team to concentrate attention on critical facts or data. Given the enormous capabilities of word processing and presentation graphics software, there is no excuse for a team not to use visual aids in its reports and presentations. On the other hand, a team should not allow their graphics' capabilities to seduce them into clouding their slides with artistic backgrounds that make the words illegible. The key to success here, as it is with words, is to know your purpose and your audience, and to use your medium appropriately.

***Write ethically!*** Designers often invest themselves in the design choices they make, in time, effort, and even values. It is, therefore, not surprising that there are temptations to present designs or other technical results in ways that not only show what is favorable, but that also suppress unfavorable data or issues. Ethical designers resist this temptation and present facts fully and accurately. This means that *all* results or test outcomes, even those that are not favorable, are presented and discussed. Ethical presentations also describe honestly and directly any limitations of a design. Further, it is also important to give full credit to others, such as authors or previous researchers, where it is due. (Remember that this discussion of the seven principles began with an acknowledgment to their originator, Thomas Pearsall, and that each chapter of the book ends with references and citations.)

## 11.2 ORAL PRESENTATIONS: TELLING A CROWD WHAT'S BEEN DONE

Most design projects call for a number of both formal and informal presentations to clients, users, and technical reviewers. Such presentations may be made before the award of a contract to do the design work, perhaps focusing on the team's ability to understand and do the job in the hope of winning the contract in a competitive procurement. During the project, the team may be called upon to present their understanding of the project (e.g., the client's needs and the artifact's functions), the alternatives under consideration and the team's plan for selecting one, or simply their progress toward completing the project. After a design alternative has been selected by the team, the team is often asked to undertake a design review before a technical audience to assess the design, identify possible problems, and suggest alternate solutions or approaches. At the end of a project, design teams usually report on the overall project to the client and to other stakeholders and interested parties.

Because of the variety of presentations and briefings that a team may be called upon to make, it is impossible to examine each of them in detail. However, there are key elements common to most of them. Foremost among these are the needs to identify the audience, outline the presentation, develop appropriate supporting materials, and practice the presentation.

### 11.2.1 Knowing the Audience: Who's Listening?

Design briefings and presentations are given to many types of audiences. Consider the new beverage container whose design we began in Chapter 3. Our design work might have to be presented to logistics managers who are concerned with how the containers will be shipped to warehouses around the country. The marketing department, concerned with establishing brand identity with the design, might want to hear about our design alternatives. Similarly, manufacturing managers will want to be briefed about any special production needs. Thus, a team planning a briefing should consider factors such as varying levels of interest, understanding, and technical skill, as well as the available time. We can assume that most attendees at a meeting are interested in at least some aspect of a project, but it is generally true that most are only interested in particular dimensions of that project. A team usually can identify such interests and other dimensions simply by asking the organizer of the meeting.

Once the audience has been identified, a team can tailor its presentation to that audience. As with other deliverables, the presentation must be properly organized and structured: The first step is to articulate a rough outline; the second is to formulate a detailed outline; and the third is to prepare the proper supporting materials, such as visual aids or physical models.

### 11.2.2 The Presentation Outline

Just as with a report, a presentation must have a clear structure. We achieve this structure by developing a rough outline. This presentation structure and organization, which should be logical and understandable, guides the preparation of supporting dialogue and discussion. And because a design presentation is neither a movie nor a novel, it should not have a "surprise ending." A sample presentation outline would include the following elements:

- *A title slide* that identifies the client(s), the project, and the design team or organization responsible for the work being presented. This slide should include company logos.
- *A roadmap for the presentation* that shows the audience the direction that the presentation will take. This can take the form of an outline, a flowchart, a big picture slide, and so on.
- *A problem statement*, which includes highlights of the revised problem statement that the team produced after research and consultation with the client.
- *Background material on the problem*, including relevant prior work and other materials developed through team research. References should be included but may be placed in a slide at the end of the presentation.
- *The key objectives of the client and users* as reflected in the top level or two of the objectives tree.
- *The key constraints* that the design must meet.
- *Functions that the design must perform*, focusing on basic functions, and means for achieving those functions.
- *Design alternatives*, particularly those that were considered at the evaluation stage, including diagrams and descriptions of each.
- *Highlights of the evaluation procedure and outcomes*, including key metrics or objectives that bear heavily on the outcome.
- *The selected design*, explaining why this design was chosen.
- *Features of the design*, highlighting aspects that make it superior to other alternatives and any novel or unique features.
- *Proof-of-concept testing*, especially for an audience of technical professionals for whom this is likely to be of great interest.

- *A demonstration of the prototype*, assuming that a prototype was developed and that it can be shown. Video or still photos may also be appropriate here.
- *Conclusion(s)*, including the identification of any future work that remains to be done, or suggested improvements to the design.

There may not always be enough time to include all of these elements in a talk or presentation, so we may need to exclude some of them. This decision will also depend at least in part on the nature of the audience.

Once the rough outline has been articulated, we should also develop a detailed outline of the presentation. This is important to ensure that everyone on the team understands every point being made, throughout the presentation. A presentation outline also helps us develop corresponding bullets or similar entries in our slides because slide bullets usually correspond to entries in the detailed outline.

Preparing a detailed outline for the presentation may seem like a great deal of work. Team members with public speaking experience may be resistant to such tasks, most likely because they have already internalized a similar method of preparation. However, since presentations represent the entire team, every member of a team should review the structure and details of their presentations, as well as the detailed outline required by such reviews.

### 11.2.3 Presentations are Visual Events

Just as a team needs to know its audience, it should also try to know the setting in which the presentation will be made. Some rooms will support certain types of visual aids, while others will not. At the earliest stages of the presentation planning, the design team should find out what devices (e.g., overhead projectors, computer connections, projectors, and whiteboards) are available and the general setting of the room in which it will be presenting. This includes its size and capacity, lighting, seating, and other factors. Even if a particular device or setup is said to be available, it is always wise to bring along backups (e.g., files on drives, transparencies, hard copies) to back up a slide presentation.

There are other tips and pointers to keep in mind about visual aids, including:

- Limit the number of slides. A reasonable estimate of the rate of slides at which slides can be covered is 1–2 slides per minute. If too many slides are planned, the presenter(s) will end up rushing through the slides in the hope of finishing. This makes for a far worse talk than a smaller selection wisely used.
- Be sure to introduce yourself and your teammates on the title slide. This is also an appropriate time for a brief overall description of the project and acknowledgment of the client. Inexperienced speakers often have the tendency to flash the title slide and move on, instead of using it as an opportunity to introduce the project and people involved.
- Beware of "clutter." Slides should be used to highlight key points; they are not a direct substitute for the reasoning of the final report. The speaker should be able to expand upon the points in the slides.
- Make points clearly, directly, and simply. Slides that are too flashy or clever tend to detract from a presentation.

- Use color skillfully. Current computer-based packages support many colors and fonts, but their defaults are often quite appropriate. Also, avoid clashing colors in such professional presentations, and be sure to keep in mind that some color combinations are hard to read for audience members who are colorblind.
- Use animation appropriately. An animated video of the function of your design might be very informative, while text flying in from the edges of the slide may not.
- Do not reproduce completed design tools (e.g., objectives trees, large morph charts) to describe the outcomes of the design process as they will likely be far too small to read. Instead, highlight selected points of the outcomes and refer the audience to a report for more detailed information.
- Consider carefully the size and distance of the audience if images of design drawings are being shown. Many line drawings are hard to display and often harder still to see and interpret in large rooms.

Remember that audiences tend to read visual aids as a speaker is talking, so he does not need to read or quote those slides. The visual aids can be simpler (and more elegant) in their content because they are there to reinforce the speaker, rather than the other way around.

### 11.2.4 Practice Makes Perfect, Maybe . . .

Presenters and speechmakers are usually effective because they have extensive experience. They have given many speeches and made many presentations, as a result of which they have identified styles and approaches that work well for them. Design teams cannot conjure up or create such real-world experience, but they can practice a presentation often enough to gain some of the confidence that experience breeds. To be effective, speakers typically need to practice their parts in a presentation alone, then in front of others, including before an audience with at least some people who are not familiar with the topic.

Another important element of effective presentation is to use words and phrases that are natural to the speaker. Each of us normally has an everyday manner of speech with which we are comfortable. While developing a speaking style, however, we have to keep in mind that ultimately we want to speak *to* an audience in *their* language, and that we want to maintain a professional tone. Thus, when practicing alone, it is useful for a presenter to try saying the key points in several different ways as a means of identifying and adopting new speech patterns. Then, as we find some new styles that work, we should repeat them often enough to feel some ownership.

Practice sessions, whether solitary or with others, should be timed and done under conditions that come as close as possible to the actual environment. Inexperienced speakers typically have unrealistic views of how long their talk will last, and they also have trouble setting the right pace, going either too fast or too slow. Thus, timing the presentation—even setting a clock in front of the presenter—can be very helpful. If slides (or transparencies or a computer) are to be used in the actual presentation, then slides (or transparencies or a computer) should be used in practice.

The team should decide in advance how to handle questions that may arise. This should be discussed with the client or the sponsor of the presentation before the team has finished practicing. There are several options available for handling questions that are asked during a talk, including deferring them to the end of the talk, answering them as they arise, or limiting questions during the presentation to clarifications of facts while deferring others until later. The nature of the presentation and the audience will determine which of these is most appropriate, but the audience should be told of that choice at the start of the presentation. When responding to questions, it is often useful for a speaker to repeat the question, particularly when there is a large audience present or if the question is unclear. The presenter or the team leader should refer questions to the appropriate team members for answers. If a question is unclear, the team should seek to clarify it before trying to answer it. And as with the presentation itself, the team should practice handling questions that it thinks might arise. It is important to have a strategy for handling questions such that the team is not talking over each other or correcting each other while answering questions.

While practicing its presentation, a team ought to prepare for questions from its audience by:

- generating a list of questions that might arise, and their answers;
- preparing supporting materials for points that are likely to arise (e.g., backup slides that may include computer results, statistical charts, and other data that may be needed to answer anticipated questions); and
- preparing to say "I don't know," or "We didn't consider that." This is very important: A team, that is, to be caught *pretending* to know has undermined its credibility and invited severe embarrassment.

A final note about selecting speakers is in order. Depending on the nature of the presentation and the project, a team may want to have all members speak (for example, to meet a course requirement); it may want to encourage less experienced members to speak in order to gain experience and confidence; or it may want to tap its most skilled and confident members. As with so many of the presentation decisions, choosing a "batting order" will depend on the circumstances surrounding the presentation. This means that, as with all of the other matters we've touched upon, a team should carefully consider and consciously decide its speaking order.

### 11.2.5 Design Reviews

A design review is a unique type of presentation, quite different from all the others that a design team is likely to do. It is also particularly challenging and useful to the team. As such, a few points about design reviews are worth noting.

A design review is typically a long meeting at which the team presents its design choices in detail to an audience of technical professionals who are there to assess the design, raise questions, and offer suggestions. The review is intended to be a full and frank exploration of the design, and it should expose the implications of solving the design problem at hand or even of creating new ones. A typical design review will consist of a briefing by the team on the nature of the problem being addressed, followed by an extensive presentation of the proposed solution. In the cases of artifacts, the team will often present an organized set of drawings or sketches that allows its audience to understand and question the team's design choices. In some cases, these materials may be provided to the attendees in advance.

A design review is often the best opportunity that the team will have to get the undivided attention of professionals about their design project. It is also worrisome for the design team, because its members may be asked to defend their design and answer pointed questions. A design review thus offers both a challenge and an opportunity to the team, giving it a chance to display its technical knowledge and its skills in constructive conflict. Questions and technical issues should be fully explored in a positive, frank environment. To benefit from the design review, the team should try to resist the natural defensiveness that comes from having its work questioned and challenged. In many cases the team can answer the questions raised, but sometimes they cannot. Depending on the nature of the meeting, the team may call upon the expertise of all of the participants to suggest new ways to frame the problem or even the design itself.

Not surprisingly, such reviews can last several hours, or even a day or two. One important decision for the team is to determine, during the review, when a matter has been adequately covered and move on. This is a real challenge, since there is a natural temptation to move on quickly if the discussion suggests that a design must be changed in ways the team doesn't like. There may be a similar temptation if the team feels that the review participants have not really "heard" the team's point of view. It is important to resist both urges: Time management should not become a cover for hiding from criticism or belaboring points.

A final point about design reviews is the need to remember that conflict in the realm of ideas is generally constructive, while personality-oriented criticism is destructive. Given the heat and light that sometimes arise at design reviews, team leaders and team members (as well as the members of the audience) must continually maintain the review's focus on the design, and not on the designers.

## 11.3 THE PROJECT REPORT: WRITING FOR THE CLIENT, NOT FOR HISTORY

The usual purpose of a final or project report is to communicate with the client in terms that ensure the client's thoughtful acceptance of a team's design choices. The client's interests demand a clear presentation of the design problem, including analyses of the needs to be met, the alternatives considered, the bases on which decisions were made, and, of course, the decisions that were taken. The results should be summarized in clear, understandable language. Highly detailed or technical materials are often placed in appendices at the end of the report, in order to support clarity. In fact, it is not unusual (and in large public works projects it is the norm) for all of the technical and other supporting materials to be moved to separate volumes. This is especially important when the client and the principal stakeholders are not engineers or technical managers, but perhaps members of the general public.

The process of writing a final report, like so much of design, is best managed and controlled with a structured approach. The design process and report writing are strikingly similar, especially in their early, conceptual stages. As with the design process, structure is not intended to displace initiative or creativity. Rather, we find that structure helps us learn how to create an organized report of our design results. One structured process that a design team might follow would include the following steps:

- determine the purpose and audience of the technical report;
- construct a rough outline of the overall structure of the report;

- review that outline within the team and with the team's managers or, in case of an academic project, with the faculty advisor;
- construct a topic sentence outline (TSO) and review it within the team;
- distribute individual writing assignments and assemble, write, and edit an initial draft;
- solicit reviews of the initial draft from managers and advisors;
- revise and rewrite the initial draft to respond to the reviews; and
- prepare the final version of the report and present it to the client.

We now discuss these steps in greater detail.

### 11.3.1 The Purpose of and Audience for the Final Report

We have already discussed determining the purpose and audience of the report in general terms. Several points should be noted in the case of a final report. The first is that the report is likely to be read by a much wider audience than simply the client's liaison with whom the team has been interacting. In this respect, the team needs to determine whether or not the liaison's interests and levels of technical knowledge are representative of the audience for the final report. The liaison may be able to guide the team to a better understanding of the expected readers, and may highlight issues of particular concern.

Another important element here is for the team to understand what the report's recipient hopes to do with the information in the final report. If, for example, the intent of the project was to create a large number of conceptual design alternatives, the audience is likely to want to see a full presentation of the design space that was explored. If, on the other hand, the client simply wanted a solution to a particular problem, they are much more likely to want to see how well the selected alternative meets the specified need.

A project report often has several different audiences, in which case the team will have to organize information to satisfy each of these target groups. This may include using technical supplements or appendices, or it may call for a structure that begins with general language and concepts, and then explores these concepts in technical subsections. The team, however, should write clearly and well for each audience, regardless of the organizational principle selected.

### 11.3.2 The Rough Outline: Structuring the Final Report

It would be foolish to start building a house or an office building without first analyzing the structure being built and organizing a construction process. Yet many people sit down to prepare a technical report and immediately begin writing, without trying to lay out in advance all of the ideas and issues that need to addressed, and without considering how these ideas and issues relate to each other. One result of such unplanned report writing is that the report turns into a project history, or worse, it sounds like a "What I Did Last Summer" essay: First we talked to the client, then we went to the library, then we did research, then we did tests, and so on. While technical reports may not be as complex as high-rise buildings or airplanes, they are nonetheless too complicated to be written as simple narratives or chronologies. Reports must be planned.

The first step in writing a good project report is building a rough outline that lays out the report's overall structure. That is, we identify the major sections into which the report is divided, which typically are:

- abstract;
- executive summary;
- introduction and overview;
- problem statement and problem definition or framing, including relevant prior work or research;
- design alternatives considered;
- evaluation of design alternatives and basis for design selection;
- results of the alternatives analysis and design selection;
- supporting materials, often set out in appendices, including;
- drawings and details;
- fabrication specifications;
- supporting calculations or modeling results; and
- other materials that the client may require.

This outline looks like a table of contents, as it should, because a final report of an engineering or design project must be organized so that a reader can go to any particular section and see it as a clear and coherent standalone document. It is not that we think things should be taken out of context. Rather, it is that we expect each major section of a report to make sense all by itself; that is, it should tell a complete story about some aspect of the design project and its results.

When should we prepare a rough outline? Indeed, when should we write our final report? It is evident that we can't write a *final* report until we have completed our work and identified and articulated a final design. However, it can be very helpful to develop a general structure for the final report early in project. We can then track and appropriately file or label key documents from the project (e.g., research memoranda, drawings, and objectives trees) according to where and if their contents would appear in the final report. Thinking about the report early on also emphasizes thinking about a project's *deliverables*, that is, those items that a team is chartered or contracted to deliver to its client during the project. We may find that the final stages or endgame of our project are much less stressful if we had organized our final report early on, simply because there will be fewer last-minute things to identify, create, and edit for insertion into the final report.

### 11.3.3 The Topic Sentence Outline: Every Entry Represents a Paragraph

A cardinal rule of writing states that *every single paragraph* of a piece should have a topic sentence that indicates that paragraph's intent or thesis. Once the rough outline of a report has been established, it is usually quite useful to build a corresponding, detailed *topic sentence outline* (TSO) that identifies the themes or topics that, collectively, make up the report. Thus, if a topic is identified by an entry in the TSO, we can assume that there is a paragraph in which that topic is covered.

The TSO enables us to follow the logic of the argument or story and assess the completeness of each section being drafted, as well as of the report as a whole. Suppose there is only one entry in a TSO for something that we consider important, say, the evaluation of alternatives. One implication of this is that the final report will have only one paragraph devoted to this topic. Since the evaluation of alternatives is a central issue in design, it is quite likely that there should be entries on a number of aspects, including the evaluation metrics and methods, the results of the evaluation, key insights learned from the evaluation, the interpretation of numerical results (especially for closely rated alternatives), and the outcome of the process. Thus, a quick examination of a TSO shows us that a proposed report is not going to address all of the issues that it should.

For the same reason, TSOs help identify appropriate cross-references that should be made between subsections and sections as different aspects of the same idea or issue are addressed in different contexts. The format of a TSO also makes it easier to eliminate needless duplication because it is much easier to spot repeated topics or ideas. In Section 11.4 we will show examples that demonstrate some of these points.

It is hard to write this way, but TSOs provide a number of advantages to a design team. First, a TSO forces the team to agree on the topics to be covered in each section. It quickly becomes clear if a section is too short for the material, or if one of the coauthors (or team members) is "poaching" on another section that was agreed upon in the rough outline. Second, a good TSO makes easier for team members to take over for one another if something comes up to prevent a "designated writer" from actually writing. For example, a team member may suddenly find that the prototype is not working as planned and she needs to do some more work on it. TSOs also make life easier for the team's report editor (see the next section) to begin to develop and use a single voice.

Finally, notwithstanding our definition of the abbreviation TSO, the entries in a TSO do not really have to be grammatically complete sentences. However, they should be complete enough that their content is clear and unambiguous.

### 11.3.4 The First Draft: Turning Several Voices Into One

One advantage of a rough outline and a topic sentence outline is that their structure allows teams members to write in parallel or simultaneously. However, this advantage comes at a price, most notably that of corralling the efforts of several writers into a single, clear, coherent document. Simply put, the more writers, the greater the need for a single, authoritative editor. Thus, one member of the team should enjoy the rights, privileges, and *responsibilities* pertaining to being the editor. Further, the team should designate an editor as soon as the planning of the report begins, hopefully at or near the onset of the project.

The editor's role is to ensure that the report flows continuously, is consistent and accurate, and speaks in a single voice. *Continuity* means that topics and sections follow a logical sequence that reflects the structure of the ideas in the rough outline and the TSO. *Consistency* means that the report uses common terminology, abbreviations and acronyms, notation, units, similar reasoning styles, and so on, throughout the report and all of its appendices. It also means, for example, that the team's objectives tree, pairwise comparison chart, and evaluation matrix all have same elements; if not, discrepancies should be noted explicitly and explained.

*Accuracy* requires that calculations, experiments, measurements, or other technical work are done and reported to appropriate professional standards and current best practices. Such standards and practices are often specified in contracts between a design team and its client(s). They typically provide that stated results and conclusions must be supported by the team's prior work. Accuracy, as well as intellectual honesty, also requires that technical reports do not make unsupported claims. There is often a temptation in a project's final moments to add to a final report something that wasn't really done well or completely.

The *voice* or style of a report reflects the way in which a report "speaks" to the reader, in ways very similar to how people literally speak to each other. It is essential that a technical report *speaks with a single voice*—and ensuring that single voice is one of the editor's most important duties. This mandate has several facets, the first of which is that the report has to read (or "sound") as if it was written by one person, even when its sections were written by members of a very large team. The president of the United States sounds like the same, familiar person, even while using several speechwriters. Similarly, a technical report must read in a single voice. Further, that voice should normally be more formal and impersonal than the voice of this book. Technical reports are not personal documents, so they should not sound overly familiar or idiosyncratic. And it is important that the voice of the report be the same, from the opening abstract through the closing conclusions, and to the last appendix.

Clearly, there are serious issues for the team dynamics of the writing process. Team members have to be comfortable surrendering control of pieces they have written, and they have to be willing to let the editor do her job. We will discuss aspects of the team dynamics of report writing in Chapter 15.

### 11.3.5 The Final, Final Report: Ready for Prime Time

A good review process ensures that a draft final report gets thoughtful reconsideration and meaningful revision. Draft reports benefit from careful readings and reviews by team members, managers, client representatives or liaisons, faculty advisors, as well as by people who have no connection with the project. This means that as we are trying to wrap up our project report, we need to incorporate reviewers' suggestions into a final, high-quality document. There are a few more points to keep in mind.

A final report should be professionally done and *polished*. This does not mean that it needs glossy covers, fancy type and graphics, and an expensive binding. Instead, it means that the report is clearly organized, easy to read and understand, and that its graphics or figures are also clear and easily interpreted. The report should also be of reproducible quality because it is quite likely to be photocopied and distributed within the client's organization, as well as to other individuals, groups, or agencies.

We should also keep in mind that a report may go to a very diverse audience, not simply to peers. Thus, while the editor needs to ensure that the report speaks with a single voice to an anticipated audience, she should try as much as possible to also ensure that the report can be read and understood by readers who may have different skill levels or backgrounds than either the design team or its client. In addition, an *executive summary* is one way to address readers who may not have the time or interest to read all of the details of the entire project.

Finally, the final report will be read and used by client(s), who will, one hopes, adopt the team's design. This means that the report, including appendices and supporting materials, is sufficiently detailed and complete to stand alone as the final documentation of the work done.

## 11.4 FINAL REPORT ELEMENTS FOR THE DANBURY ARM SUPPORT

As required in most design projects, the student teams responsible for the arm support design for the Danbury School reported their results in the form of final reports and oral presentations. In this section we look briefly at some of the intermediate work products associated with their reports to gain further insight into some of the "do's and don'ts" discussed in Section 11.3.

### 11.4.1 Rough Outlines of Two Project Reports

The two teams we have followed each prepared a rough outline as a first step in laying out the report structure. Tables 11.1 and 11.2 respectively show the rough outlines developed by teams A and B and they are similar, yet different. Team A, for example, dedicated several sections to justifying their final design, while the team B organized around process.

**TABLE 11.1 Danbury arm support team A's rough outline. The rough outline should show the overall structure of the report in a way that allows team members to divide up work with little or no unintended duplication. The structure should also proceed in a clear and logical manner. Does it for this report?**

I. Introduction
II. Description of problem definition
  a. Problem statement
  b. Design objectives and constraints
III. Generation of design alternatives
  a. Morphological chart
  b. Description of design alternatives
  c. Description of subcomponents
IV. Design selection process
  a. Metrics description
  b. Metrics application
V. Final design
  a. Detailed description
  b. Prototype details
VI. Testing the design
VII. Conclusions
  a. Strengths and weaknesses of final design
  b. Suggestions for more advanced prototype
  c. Recommendations to the client
VIII. References
Appendix: Work breakdown structure
Appendix: Pairwise comparison chart

Adapted from Attarian et al. (2007).

**TABLE 11.2 Danbury arm support team B's rough outline. As with the outline presented in Table 11.1, this outline also shows the overall structure, and it is clear that the report focuses on reporting detailed testing and evaluation of the chosen design**

Introduction

I. Problem statement
II. Background information on cerebral palsy, motivation for project
III. Design plan
  a. Work breakdown structure
  b. Definition of objectives and constraints, including objectives tree
  c. Definition of functions and means, morphological chart
IV. Design research
  a. Summary of devices currently available
  b. Evaluation of these devices for suitability in this project
V. Description and evaluation of design alternatives
  a. Details and drawings of each alternative
  b. Metrics for choosing between designs
VI. Final design
  a. Detailed description of chosen alternative
  b. Description of prototype and how it works
VII. Testing the design
  a. Description of three test sessions at Danbury
  b. Conclusions and refinements of design based on testing
VIII. Design evaluation
  a. Consideration of constraints
  b. How well design meets objectives
  c. Functional analysis
  d. Details on proposed design changes based on testing and evaluation
IX. Works cited

Appendix: Work breakdown structure

Appendix: Research on dashpots

Adapted from Best et al. (2007).

Both teams relegated sketches and drawings to appendices, although the second team put building instructions in the report body. This reflects the freedom that teams have to decide on an appropriate structure to convey their design results. This freedom, however, does not excuse them from having a logical ordering that allows the reader to understand the nature of the problem or the benefits of their solution.

Looking at the structure of these final reports, we also see just how much of each could have been written during the course of the project. Both teams used the formal design tools discussed in Chapters 3–8 to document their decision processes. Thus, the teams could—and should—have tracked and organized their work products in order to facilitate the writing of their final reports.

Finally, neither outline will adequately translate directly into a report. There are issues that could be considered in more than one section, and others not covered at all. Unless a team follows the outline with a TSO or some other detailed plan, the first draft of its final report will need an unnecessarily high degree of editing.

## 11.4.2 A TSO for the Danbury Arm Support

Table 11.3 shows an excerpt from the topic sentence outline prepared by team B. Notice that while each entry is not in itself a complete sentence, the specific point of that entry is easy to see. At this level of detail, it is relatively easy to identify points that are either redundant or inadequately covered.

The TSO enables the team to see not only what will be covered within each section, but also within each paragraph of the section. It also permits team members to take issue with or make suggestions about a section before writing and "wordsmithing" efforts are extended. For example, team B's definitions of metrics are not clear and could be challenged by a nit-picking reader (such as a professor or a technical manager). It is also clear that the ideas conveyed in some single paragraphs might be better explained were they divided into two paragraphs. It is also not clear why attention was specifically called to the liaison's definitions of constraints: Didn't the constraints (and objectives and revised problem statement) emerge out of the team's questioning of and discussions with the client? Elsewhere in the report the team implies that it alone developed the list of objectives. Perhaps this TSO indicates confusion about how the design process was executed, or at least how its execution is to be reported.

**TABLE 11.3 An excerpt from the TSO for one section of team B's final report outline (shown in Table 11.2) for the Danbury arm support project**

III. Design plan

- A. After clarifying the problem statement, the team began the process of designing the device
  - a. Paragraph describing the overall approach to the design
    - i. Work breakdown structure
    - ii. Objectives and constraints
    - iii. Defining functions and means
    - iv. Creating and evaluating design alternative
- B. The work breakdown structure consists of the tasks for the design process and their deadlines
- C. In order to implement the design, the team needed to define objectives and constraints
  - a. Paragraph defining objectives
    - i. Objectives are things one wants the design to achieve
    - ii. Objectives have a hierarchy
    - iii. List of ranked objectives
  - b. Paragraph on liaison reaction to ranked objectives list
    - i. Liaisons added an objective and ranked it
  - c. Paragraph on organization of objectives
    - i. Objectives sorted into three categories: user-friendly, primary device functions, and features
    - ii. Objectives divided into subobjectives
    - iii. Listing of subobjectives
    - iv. Objectives tree
  - d. Paragraph on evaluating objectives using metrics
  - e. Paragraph defining constraints
    - i. Constraints are limits on the design
    - ii. List of constraints and description
  - f. Paragraph on liaison input and reaction to constraints
    - i. Initial constraints
    - ii. Constraints added after reaction from liaisons

Adapted from Best et al. (2007).

We also note that team B has cast their process as a historical narrative, which is very much at odds with our recommended style. For example, the statement "After clarifying the problem statement, the team began the process of designing the device . . . " is a red flag that the team is documenting the passing of time and events, not the evolution of a design process. Fortunately, because effort was invested in a TSO, it was relatively easy to identify needed changes and to make them.

### 11.4.3 The Final Outcome: The Danbury Arm Support

The teams we have followed working on the Danbury arm support project did finally finish their work. They produced formal oral presentations, prototypes, and final reports. The two team's reports were as different as were their designs. For instance, the reports were, respectively, 18 and 61 pages long! We have also shown some of their drawings in Figures 9.3 and 9.7, team B's model in Figure 10.1, and team A's prototype in Figure 10.2.

While the two designs have obvious similarities, they also have clear differences. For example, their mounting structures differ, as do their use of damping devices. These differences and similarities are not a surprise to designers or to engineering faculty, and they should not surprise you. As we have also noted previously, design is an *open-ended* activity, that is, there is no single—or even assured—solution to a design problem. Dealing with the uncertainty implied by the absence of a guaranteed, unique outcome is the reason that design is both challenging and exciting. The only thing that is certain is the satisfaction and excitement experienced by designers, clients, and users when a good design is achieved.

## 11.5 NOTES

*Section 11.1*: As noted in the text, the seven principles of technical writing are drawn from Pearsall (2001). In addition to Pearsall, there are a number of excellent books to support technical writing, including Pfeiffer (2001), Stevenson and Whitmore (2002), and the classic Turabian (1996). There is no better reference to effective use of graphics than Tufte (2001), a classic which belongs in the library of every engineer.

*Section 11.4*: The final results for the Danbury arm support design project from Attarian et al. (2007) and Best et al. (2007).

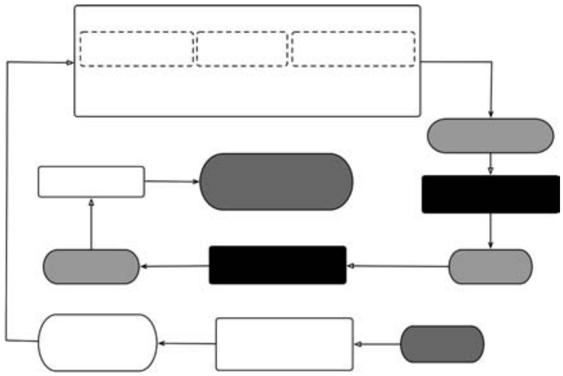
# PART IV

# *DESIGN MODELING, ENGINEERING ECONOMICS, AND DESIGN USE*

CHAPTER 12

# MATHEMATICAL MODELING IN DESIGN

*Math and physics are very much part of the design process!*



**MATHEMATICAL MODELS** are central to design because we have to be able to *predict* the behavior of the devices or systems that we are designing. Every new airplane or building, for example, represents a model-based prediction that the plane will fly or the building will stand without producing unintended, often tragic, consequences. It is important for us to ask: How do we create mathematical models? How do we validate such models? How do we use them? And, are there any limits on their use?

We can't possibly introduce here all of the models and techniques needed to model all the kinds of designs that engineers do. However, we can illustrate some major points and "habits of thought" by analyzing some very basic mechanical and electrical devices. In particular, after discussing some fundamental mathematical modeling ideas, we will model a basic circuit found in electrically powered toy cars and then analyze and design a ladder rung.

## 12.1 SOME MATHEMATICAL HABITS OF THOUGHT FOR DESIGN MODELING

If a client wants a device that has reduced energy consumption as an objective (or a limit on energy consumption as a constraint), we need to have a model of our design to address our client's concerns. Similarly, if we know that users of a ladder will be unsettled by a deflection of a rung of more than 0.5 in., we need a model of the bending of the rung to figure out how to meet that as an objective (i.e., the movement should be small) or as a constraint (e.g., specifying a maximum value). When we talk about models and modeling, we use *model* both as a *verb* to denote the activity in which we think about and make representations of how devices or objects of interest behave; and as a *noun* to denote the representation of how devices or objects of interest behave. Those representations can be, in principle, in words, drawings or sketches, physical models, computer programs, or mathematical formulations. For our present purposes, we will focus on *representing the behavior and function of real devices in mathematical terms*.

### 12.1.1 Basic Principles of Mathematical Modeling

Mathematical modeling is an activity with underlying principles and a host of methods and tools. The overarching principles are almost philosophical in nature:

- ***Why*** do we need a model?
- For what will we ***use*** the model?
- What do we want to ***find*** with this model?
- What data are we ***given***?
- What can we ***assume***?
- ***How*** should we develop this model, that is, what are the appropriate physical principles we need to apply?
- What will our model ***predict***?
- Can we ***verify*** the model's predictions (i.e., are our calculations correct?)
- Are the predictions ***valid*** (i.e., do our predictions conform to what we observe?)
- Can we ***improve*** the model?

This list of questions is *not* an algorithm for mathematical model building. The underlying ideas are key to problem formulation generally. Thus, the individual questions will recur often during the modeling process, and the list should be regarded as a general approach to *habits of thought* for mathematical modeling, which *are essential for good design modeling*.

### 12.1.2 Abstractions, Scaling, and Lumped Elements

An important decision in modeling is choosing the right level of detail for the problem, which thus dictates the level of detail for the model. We call this part of the modeling process *abstraction*. It requires a thoughtful approach to identifying the phenomena to be emphasized, that is, to answering the fundamental question about *why* a model is being developed and how we intend to *use* it. Stated differently, thinking about finding the right level of abstraction or detail means identifying the right *scale* for our model means thinking about the magnitude or size of quantities measured with respect to a standard that has the same physical dimensions.

For example, a linear elastic spring, usually expressed in terms of Hooke's law, $F = kx$, can be used to model more than just the relation between force and relative extension of a simple coiled spring. It can be used to describe the static load-deflection behavior of a ladder's step, and the spring constant $k$ will reflect the stiffness of the step taken as a whole. This interpretation of $k$ incorporates detailed properties of the step, such as the material of which it is made and its dimensions. That same spring equation can also used to model how tall buildings respond to wind loading and to earthquakes. Both of these examples suggest that we can build a simple model of a building by aggregating various details within the parameters of that model. That is, we would compile or *lump* together a lot of information about how the building is framed, its geometry, its materials, and so on, into a stiffness $k$ for the ladder step or the building. We need detailed expressions for each device that allow us to aggregate their particular properties into the overarching step or building model. And, of course, we would *validate* our linear spring models by observing and measuring the respective behaviors of the ladder step or the tall building.

We can also use springs to model atomic bonds if we can develop or show how their spring constants depend on atomic interaction forces, atomic distances, subatomic particle dimensions, and so on. Thus, we can use a linear spring at both very small, *micro* scales to model atomic bonds, and at very large *macro* scales, as for buildings. The notion of scaling includes several ideas, including the effects of geometry on scale, the relationship of function to scale, and the role of size in determining limits—all of which we need to choose the right scale for a model in relation to the "reality" we want to capture.

Going a step further, we often say that a "real," three-dimensional object behaves like a simple spring. When we say this, we are introducing the idea of a *lumped element model* in which the actual physical properties of a real object or device are aggregated or *lumped* into less detailed, more abstract expressions. For example, we can model an airplane in very different ways, depending on our goals. To lay out a flight plan or trajectory, we can simply consider the airplane as a point mass moving with respect to a spherical coordinate system: The total mass of the plane is lumped into a point mass; the effect of the surrounding atmosphere is modeled by introducing a retarding drag force to act on the mass point in some proportion to the mass' relative speed. To model more local effects of air moving around the plane's wings, our model would have to account for the wing's shape and surface area and be sufficiently complex to incorporate the aerodynamics of different flight regimes. To model (and design) the flaps used to control the plane's ascent and descent, our model would have to include a system to control the flaps and account for the wing's strength and vibration response. Again, what we lump into our lumped elements depends on the scale on which we choose to model, which depends in turn on our intentions for that model.

## 12.2 SOME MATHEMATICAL TOOLS FOR DESIGN MODELING

We now present some tools that we can use to apply the "big picture" principles to develop, use, verify, and validate mathematical models. These tools include dimensional analysis, approximations of mathematical functions, linearity, and conservation and balance laws.

### 12.2.1 Physical Dimensions in Design (I): Dimensions and Units

One central idea in mathematical modeling is the following: Every independent term in every equation we use has to be *dimensionally homogeneous* or *dimensionally consistent*, that is, every term has to have the same net physical dimensions. Thus, every term in a balance of mass must have the dimension of mass, and every term in a summation of forces must have the physical dimension of force. We also call dimensionally consistent equations *rational* equations. In fact, one important way of validating newly developed mathematical models (or of confirming formulas before using them for calculations) is to ensure that they are rational equations.

The physical quantities used to model objects or systems represent *concepts*, such as time, length, and mass, to which we attach *numerical* measurements or values. If we say a soccer field is 60 m wide, we are invoking the concept of length or distance, and our numerical measure is 60 m. The numerical measure implies a comparison with a standard or scale: Common measures provide a frame of reference for making comparisons.

We define two classes for the physical quantities we use to model problems, fundamental and derived:

- *Fundamental* or *primary* quantities can be measured on a scale that is independent of those chosen for any other fundamental quantities. In mechanical problems, for example, mass, length, and time are usually taken as the fundamental mechanical *dimensions* or variables.
- *Derived* quantities generally follow from definitions or physical laws, and they are expressed in terms of the dimensions that were chosen as fundamental. Thus, force is a derived quantity that is defined by Newton's law of motion.

If mass, length, and time are chosen as primary quantities, then the dimensions of force are (mass $\times$ length)/(time)$^2$. We use the notation of brackets [ ] to read as "the dimensions of." If M, L, and T stand for mass, length, and time, respectively, then

$$[F = \text{force}] = (\mathrm{M} \times \mathrm{L})/(\mathrm{T})^2 \tag{12.1}$$

Similarly, $[A = \text{area}] = (\mathrm{L})^2$ and $[\rho = \text{density}] = \mathrm{M}/(\mathrm{L})^3$. Also, for any given problem, we have to have enough fundamental quantities to be able to express each derived quantity in terms of those primary quantities.

The *units* of a quantity are the numerical aspects of a quantity's dimensions expressed in terms of a given physical standard. Thus, a unit is an arbitrary multiple or fraction of a physical standard. The most widely accepted international standard for measuring length is the meter (m), but length can also be measured in units of centimeters (1 cm = 0.01 m) or of feet (0.3049 m). The magnitude or size of the attached number obviously depends on the unit chosen, and this dependence often suggests a choice of units to facilitate calculation or communication. For example, a soccer field width can be said to be 60 m, 6000 cm, or (approximately) 197 ft.

We often want to compute particular numerical measures in different sets of units. Since the physical dimensions of a quantity are constant, there must exist numerical relationships between the different systems of units used to measure the amounts of that quantity (e.g., 1 foot (ft) = 30.48 centimeters (cm), and 1 hour (h) = 60 minutes (min) = 3600 seconds (sec *or* s)). This equality of units for a given dimension allows units to be changed or converted with a straightforward calculation. For example, units of pressure in the American system (psi) can be converted to units of pressure in the SI system (pascal):

$$1\,\frac{\text{lb}}{\text{in.}^2} \cong 1\,\frac{\text{lb}}{\text{in.}^2} \times 4.45\,\frac{\text{N}}{\text{lb}} \times \left(\frac{\text{in.}}{0.0254\ \text{m}}\right)^2 \cong 6897\,\frac{\text{N}}{\text{m}^2} \equiv 6897\ \text{Pa} \tag{12.2}$$

Each of the multipliers in this conversion equation has an effective value of unity because of the equivalencies of the various units, that is, 1 lb $\cong$ 4.45 N, and so on. This, in turn, follows from the fact that the numerator and denominator of each of the above multipliers have the same physical dimensions.

We noted earlier that each independent term in a rational equation has the same net dimensions. Thus, we cannot add length to area in the same equation, or mass to time, or charge to stiffness. On the other hand, we can add quantities having the same dimensions but expressed in different units (e.g., length in meters and length in feet), although we must be very careful. The fact that equations must be rational in terms of their dimensions is central to modeling because it is one of the best—and easiest—check to make to determine whether a model makes sense, has been correctly derived, or even correctly copied!

In a familiar model from mechanics, the speed of a particle, $V$, due to the acceleration of gravity, $g$, when dropped from a height, $h$, is given by

$$V = \sqrt{2gh} \tag{12.3}$$

Note that both sides of eq. (12.3) have the same net physical dimensions, that is, L/T on the left-hand side and $[(L/T^2)L]^{1/2}$ on the right. As a result, eq. (12.3) is *dimensionally homogeneous* because it is totally independent of the system of units being used to measure $V$, $g$, and $h$. However, we often create unit-dependent versions of such equations because they are easier to remember or they make repeated calculations convenient. For example, we may be working entirely in metric units, in which case $g = 9.8\ \text{m/s}^2$, so that

$$V(\text{m/s}) = \sqrt{2(9.8)h} \cong 4.43\sqrt{h} \tag{12.4}$$

Equation (12.4) is valid *only* when the particle's height is measured in meters. If we were working with American units only, then $g = 32.17\ \text{ft/sec}^2$ and

$$V(\text{ft/sec}) = \sqrt{2(32.17)h} \cong 8.02\sqrt{h} \tag{12.5}$$

Equation (12.5) is valid *only* when we measure the particle's height in feet. Neither eq. (12.4) nor eq. (12.5) is dimensionally homogeneous. While these formulas may be easier to remember or use, we must keep in mind their limited validity.

There is one other way in which these dimensional considerations come into play that is worth noting. In Chapter 4, we introduced sets of *units of interest* for the scales of metrics to be used for assessing the achievement of objectives. These are also often called *figures of merit*. Similarly, to optimize a design we could construct mathematical *objective functions* that represent figures of merit and whose value is to be optimized. It is very important to remember that such objective functions, like equations, should likewise be rational functions: All of the independent terms in an objective function must have the same net dimensions.

### 12.2.2 Physical Dimensions in Design (II): Significant Figures

We use numbers a lot in engineering for both design and analysis, but we often need to remind ourselves about the significance of each of those numbers. In particular, people often ask how many decimal places they are expected to keep. But that's the wrong question to ask, because the *number of significant figures* (NSF) is *not* determined by the placement of the decimal point. In scientific notation, the *number of significant figures* is equal to the number of digits counted *from* the first nonzero digit on the left *to* either (a) the last nonzero digit on the right if there is no decimal point, or (b) the last digit (zero or nonzero) on the right when there is a decimal point. (See the examples shown in Table 12.1.) This notation or convention assumes that terminal zeroes without decimal points to the right signify only the magnitude or power of 10. In fact, confusion over the NSF arises because of the presence of terminal zeros: We don't know whether those zeroes are intended to signify something, or whether they are placeholders to fill out some arbitrary number of digits.

One way to think about the NSF is to imagine that we are running a test whose outcomes could be $A$, $B$, or $C$, and we want to know how often we see the result $A$. If $A$ occurs four times in a set of 10 tests, then we would say that $A$ occurs in 0.4 of the tests. If we got $A$ 400 times in a run of 1000 tests, we would say that $A$ was found in 0.400 of the tests—but how do we make it clear that those extra two zeroes have meaning? The answer is that we can eliminate any confusion if we write all such numbers, whether from technical calculations or experimental data, in *scientific notation*. In scientific notation we write

**TABLE 12.1 Numbers written in different forms, together with the *number of significant figures* (NSF) of each and an assessment of the NSF that can be assumed or inferred. Confusion about the NSF arises because of the meaning of the terminal zeroes is not stated**

| Measurement | Significant Figures | Assessment |
|---|---|---|
| 5415 | Four | Clear |
| 5400 | Two ($54 \times 10^2$) or three ($540 \times 10^1$) or four (5400) | Not clear |
| 54.0 | Three | Clear |
| 54.1 | Three | Clear |
| 5.41 | Three | Clear |
| 0.00541 | Three | Clear |
| $5.41 \times 10^3$ | Three | Clear |
| 0.054 | Two | Clear |
| 0.0540 | Two (0.054) or three (0.0540) | Not clear |
| 0.05 | One | Clear |

numbers as products of a “new” number that is normally in the interval 1–10 and a power of 10. Thus, numbers both large and small can be written in one of two equivalent, yet unambiguous forms:

$$\begin{aligned} 514{,}000{,}000 &= 0.514 \times 10^9 = 5.14 \times 10^8 \\ 0.000075 &= 0.75 \times 10^{-4} = 7.5 \times 10^{-5} \end{aligned}$$

Also on the subject of the NSF, we should *always* remember that *the results of any calculation or measurement cannot be any more accurate than the least accurate starting value*. We cannot generate more significant digits or numbers than the smallest number of significant digits in any of our starting data. It is far too easy to become captivated by all of the digits produced by our computers or spreadsheets, but it is really important to remember that *any calculation is only as accurate as the least accurate value we started with*.

### 12.2.3 Physical Dimensions in Design (III): Dimensional Analysis

We often find it useful to work with or even create *dimensionless* variables or numbers, which by design are intended to compare the value of a specific variable with a standard of obvious relevance. For example, hydrologists model some of the behavior of soil in terms of its *porosity*, $\eta$, which is defined as the dimensionless ratio $\eta = V_v/V_t$, where $V_v$ is the volume of voids (or interstitial spaces) in the soil and $V_t$ is the total volume of the soil being considered. We also see that this definition of porosity *normalizes* or *scales* the void volume $V_v$ against the total volume $V_t$. A similar (and more famous) example is Einstein’s formula for the relativistic mass of a particle, $m = m_0/\sqrt{1-(v/c)^2}$, in which the mass $m$ is normalized against the rest mass, $m_0$, and the particle speed is scaled against the speed of light, $c$, in the dimensionless ratio $v/c$. Note that Einstein’s formula is dimensionally homogeneous, and that the particle speed is normalized such that $0 \leq v/c \leq 1$ and the mass such that $1 \leq m/m_0 < \infty$.

We can learn a lot about some behavior by doing *dimensional analysis*, that is, by expressing that behavior in a dimensionally correct equation among certain variables or dimensional groups. The *basic method* of dimensional analysis is an informal unstructured approach for determining dimensional groupings that depends on constructing a functional equation that contains all of the relevant variables, for which we know the dimensions. We then identify the proper dimensionless groups by thoughtfully eliminating dimensions. For example, consider again the free fall of a body in a vacuum described by eq. (12.3). A more general functional expression of eq. (12.3) is

$$V = V(g, h) \tag{12.6}$$

The physical dimensions of the three variables in eq. (12.6) are, respectively, $[V] = \mathrm{L/T}$, $[g] = \mathrm{L/T^2}$, and $[h] = \mathrm{L}$. The time dimension, T, appears only in the speed and gravitational acceleration, so that dividing $V$ by the square root of $g$ allows us to eliminate time and find a quantity whose remaining dimension can be expressed entirely in terms of length, that is,

$$\left[\frac{V}{\sqrt{g}}\right] = \sqrt{L} \tag{12.7}$$

If we repeat this thoughtful elimination with regard to the length dimension, we would divide eq. (12.7) by $\sqrt{h}$, which means that

$$\left[\frac{V}{\sqrt{gh}}\right] = 1 \tag{12.8}$$

Since we have but a single dimensionless group here, it follows that

$$V = \text{constant} \times \left(\sqrt{gh}\right) \tag{12.9}$$

Equation (12.9)—found with dimensional analysis alone, invoking neither Newton's law nor any other principle of mechanics—confirms eq. (12.3). This elementary application of dimensional consistency tells us something about the power of dimensional analysis. On the other hand, we do need some physics, either theory or experiment, to define the constant in eq. (12.9) that gets us to eq. (12.3).

As a further example, suppose we want to design a step on a ladder. One model of a step's behavior would be to think of its as a fixed-ended beam under a centrally applied vertical load $P$. We will present the appropriate model in Section 12.2.1 when we use it, but let us first see whether we can identify the form we will see by applying the basic method of dimensional analysis. We show in Table 12.2 the four derived variables for this problem and their respective dimensions: The deflection $\delta$ at the center of the beam, the load $P$, the beam length $L$, the bending stiffness $EI$ of the beam (which is actually the product of its modulus of elasticity $E$ and the second moment $I$ of the beam's cross-sectional area). We also see that all of our four variables are expressed in just two physical dimensions, force (F) and length (L). We then ask, dimensionally, how does the deflection $\delta$ of the beam (or step) depend on the other four variables? That is, in analogy with our dimensional analysis of eq. (12.6), we look for the dimensionless groups embodied in the functional expression:

$$\delta = \delta(P, L, EI) \tag{12.10}$$

Since we know that the beam or step deflection has the physical dimension of length, we would eliminate the force dimension by noting that $[P/EI] = \mathrm{L}^{-2}$ and then eliminating the force dimension by dividing accordingly:

$$\delta_1 = \delta_1\left(\frac{P}{EI}, L\right) \tag{12.11}$$

**TABLE 12.2 The four quantities chosen to model the fixed-ended beam that, in turn, will be a model for a step on a ladder. *P* and *L* are chosen as fundamental, and $\delta$ and *EI* are then taken as derived**

| Derived Quantities | Dimensions |
|---|---|
| Deflection ($\delta$) | L |
| Load ($P$) | F |
| Length ($L$) | L |
| Bending stiffness ($EI$) | $\mathrm{FL}^2$ |

Note that eq. (12.11) also suggests that the beam's deflection varies with the ratio $P/EI$, which makes sense physically: increase the stiffness $EI$ with respect to the load $P$ and expect the deflection to decrease, and vice versa. Then, in dimensional terms, since $[\delta_1] = \mathrm{L}$, it follows that we need to determine a value of an exponent $\alpha$ such that

$$[\delta_1] = \left[\left(\frac{P}{EI}\right)L^{\alpha}\right] = \mathrm{L}^{-2}\mathrm{L}^{\alpha} = \mathrm{L} \tag{12.12}$$

Clearly, $\alpha = 3$ and

$$\delta = \text{constant} \times \left(\frac{PL^3}{EI}\right) \tag{12.13}$$

Equation (12.13) also confirms our physical intuition because as we make the beam or step longer, we would expect it to be more flexible, that is, to deflect more.

We can describe the basic method of dimensional analysis in a series of steps:

**a.** List all of the variables and parameters of a problem and their dimensions.

**b.** Anticipate how each variable qualitatively affects quantities of interest, that is, does an increase in a variable cause an increase or a decrease?

**c.** Identify one variable as depending on the remaining variables and parameters.

**d.** Express that dependence in a functional equation (i.e., in analogs of eqs. (12.6) and (12.10)).

**e.** Choose and then eliminate one of the primary dimensions to obtain a revised functional equation.

**f.** Repeat step (e) until a revised, *dimensionless* functional equation is found.

**g.** Review the final *dimensionless* functional equation to see whether the apparent behavior accords with the behavior anticipated in step (b).

### 12.2.4 Physical Idealizations, Mathematical Approximations, and Linearity

We generally *idealize* or approximate situations or objects so that we can model them and apply those models to find behaviors of interest. We make two kinds of idealizations, physical and mathematical, and the order in which we make them is important. Think of the basic pendulum: A known mass $m$ hangs from a string of given length $l$. *First*, we identify those elements that we believe are important to the problem. We assume the string is weightless and acts only in tension, and that gravity provides the only external force. We also assume that any wind resistance is negligible and that, for now, the pendulum's swing angle $\theta(t)$ is not limited in magnitude. Our model is (still) verbal, but we have idealized several facets of the pendulum's anticipated behavior by assuming the string is weightless and by neglecting wind resistance. Soon we will also examine the consequences of considering only small angles. But for now, we have an initial *physical idealization*.

*Second*, we translate our physical idealization into a mathematical model. We have to be careful that our mathematical models are consistent with what we have assumed in our physical idealization. We start with the basic equation of motion for the pendulum, which we assume is familiar from physics. We write it here in its exact, nonlinear form, cast in terms of the pendulum's swing angle:

$$\frac{d^2\theta(t)}{dt^2} + \left(\frac{g}{l}\right)\sin\theta(t) = 0 \tag{12.14}$$

Equation (12.14) is a nonlinear differential equation because of the sinusoid term. Such nonlinearities make it difficult to find exact, closed-form solutions. While there is, in fact, an elegant, closed-form, implicit solution for the nonlinear pendulum, we won't discuss it as that's not our focus. Rather, we want to ask the question: Can we make a linear approximation of eq. (12.14)? If so, how, and what does that mean?

Engineers typically try to build models that are, mathematically speaking, *linear* models. We do this because nonlinear problems are invariably harder to solve, and also because *linear models work extraordinarily well for many devices and behaviors of interest*. In fact, one of the most often used linear approximations is a small angle approximation. The more common form of the small angle approximation is that of $\sin\theta$ for small angles $\theta$. In this case we have

$$\sin\theta = \theta - \frac{\theta^3}{3!} + \cdots \cong \theta \tag{12.15}$$

Assuming eq. (12.15) enables us to *linearize* the classical pendulum problem, that is, turn it into a linear differential equation that is easily solved:

$$\frac{d^2\theta(t)}{dt^2} + \left(\frac{g}{l}\right)\theta(t) = 0 \tag{12.16}$$

In this context, it is also worth looking at the tension in the pendulum's string,

$$T = ml\left(\frac{d\theta(t)}{dt}\right)^2 + mg\cos\theta(t) \tag{12.17}$$

as well as the pendulum system's potential energy,

$$V(t) = mgl(1 - \cos\theta(t)) \tag{12.18}$$

Both eqs. (12.17) and (12.18) contain the term $\cos\theta(t)$ so the equation becomes, how do we approximate that term for small angles $\theta(t)$? This requires thinking about the meaning of "small" and in relation to what. For small angles we could have either

$$\cos\theta = 1 - \frac{\theta^2}{2!} + \cdots \cong 1 \tag{12.19a}$$

or

$$(1 - \cos\theta) = \left[1 - \left(1 - \frac{\theta^2}{2!} + \cdots\right)\right] \cong \frac{\theta^2}{2!} \tag{12.19b}$$

Equations (12.19a) and (12.19b) clearly present very different results, either one of which could be a suitable mathematical approximation. The "trick" is to properly understand the physical idealization that we are trying to represent, which in this case means properly identifying the scale by which we measure "small." If we were simply approximating $\cos\theta$ for small angles $\theta$, then eq. (12.19a) would be appropriate, and, we would find from eq. (12.17) that $T \cong mg$ for small angles of motion. But to approximate the potential energy of a pendulum for small angles $\theta$, we would have to use eq. (12.19b) because in this case we're comparing $(1 - \cos\theta)$ to 1, and so we would find from eq. (12.18) that $V(t) \cong mgl\theta^2(t)/2$ for small angles of motion.

Linearity shows up in other contexts. Consider *geometrically similar* objects, that is, objects whose basic geometry is essentially the same. For two right circular cylinders of radius $r$ and respective heights $h_1$ and $h_2$, the total volume in the two cylinders is

$$V_{cyl} = \pi r^2 h_1 + \pi r^2 h_2 = \pi r^2 (h_1 + h_2) \tag{12.20}$$

Equation (12.20) demonstrates that the volume is *linearly proportional* to the height of the fluid in the two cylinders. Further, we obtained the total volume using the *principle of superposition* (i.e., by adding the two volumes), which we could do because the volume $V_{cyl}$ is a *linear function* of the total height $h$. Note, however, that the volume is *not* a linear function of the radius, $r$. That is, for different radii for the two cylinders, eq. (12.20) becomes

$$V_{cyl} = \pi h_1 r_1^2 + \pi h_2 r_2^2 \tag{12.21}$$

The relationship between volume and radius is nonlinear for the cylinders, so we can't calculate the total volume just by superposing or adding the two radii. This result is emblematic of what happens when a linearized model is replaced by its (originating) nonlinear version.

### 12.2.5 Conservation and Balance Laws

Many of the mathematical models used in engineering design are statements that some property of an object or system is being conserved. For example, the motion of a body moving on an ideal, frictionless path might be analyzed by stipulating that its *energy is conserved*, that is, energy is neither created nor destroyed. Sometimes, as in modeling the population of an animal colony or the volume of a river flow, *quantities that cross a defined boundary* (whether individual animals or water volumes) *must be balanced*. That is, we have to count or measure both what goes in to and what comes out of the boundary of the domain we're watching. Such *balance* or *conservation principles* are applied to assess the effect of maintaining levels of physical attributes. Conservation and balance equations are related: Conservation laws are special cases of balance laws.

The mathematics of balance and conservation laws is straightforward. We start by (conceptually, and sometimes graphically) drawing a boundary around the device or system we are modeling. If we denote the physical attribute or property being monitored as $N(t)$ and the independent variable time as $t$, a balance law for the *temporal* or time rate of change of that property within the system boundary outlined can be written as

$$\frac{dN(t)}{dt} = n_{in}(t) + g(t) - n_{out}(t) - c(t) \tag{12.22}$$

where $n_{in}(t)$ and $n_{out}(t)$ represent the flow rates of $N(t)$ into (the *influx*) and out of (the *efflux*) the system boundary, $g(t)$ is the rate at which $N$ is generated within the boundary, and $c(t)$ is the rate at which $N$ is consumed within that boundary. Equation (12.18) is also called a *rate equation* because each term has both the meaning and dimensions of the rate of change with time of the quantity $N(t)$.

In cases where there is no generation and no consumption within the system boundary (i.e., when $g = c = 0$), the balance law in eq. (12.22) becomes a *conservation law*:

$$\frac{dN(t)}{dt} = n_{in}(t) - n_{out}(t) \tag{12.23}$$

Here, then, the rate at which $N(t)$ accumulates within the boundary is equal to the difference between the influx, $n_{in}(t)$, and the efflux, $n_{out}(t)$.

Perhaps the most familiar balance and conservation laws are those associated with Newtonian mechanics. Newton's first law, usually presented as the equation of motion, can be viewed as a balance law because it refers to a balance of forces:

$$\sum \vec{F} = m\vec{a} = \frac{d}{dt}(m\vec{v}) \tag{12.24}$$

Note, however, that eq. (12.24) also represents a conservation law because it states the conservation of momentum: If there are no net forces acting on the mass $m$, then $d(m\vec{v})/dt = 0$ and the momentum $m\vec{v}$ is conserved.

The second familiar conservation principle in Newtonian mechanics is the principle of conservation of energy:

$$E(t) = \frac{1}{2}m\left(\vec{v}\cdot\vec{v} = v^2\right) + V = E_0 \tag{12.25}$$

Here $V$ represents the particular form of potential energy for the system under consideration (e.g., $mgh$ for gravitational potential and $kx^2/2$ for a linear spring) and $E_0$ is the constant total (kinetic plus potential) energy. For a nonideal system, energy is not conserved and the result is the work–energy principle:

$$\left[\frac{1}{2}mv_2^2 - \frac{1}{2}mv_1^2\right] + V_2 - V_1 = \int_1^2 \vec{F}\cdot d\vec{s} \tag{12.26}$$

Thus, the difference in the total (kinetic and potential) energy between states 1 and 2 is equal to the work done by the forces acting on the system as they traverse the path from state 1 to state 2.

### 12.2.6 Series and Parallel Connections

Almost all designs require connections, whether as simple as a mass hanging from a rope tied to a hook, or a circuit connecting a power source to a light bulb through a switch. Such connections can generally be characterized as either series or parallel. Further, we gain tremendous insights into design behavior when we link these characterizations to appropriate balance or conservation laws. We show two examples in Figure 12.1, both involving simple linear springs. The first is a *series* connection, where the two springs each support the same load $P$, while the second is a *parallel* connection, each deflecting the same total amount. In the series connection, we first write each spring's *constitutive law*, Hooke's law, as

$$F_{s1} = k_1(x_{s1} - x_{s0}) \quad \text{and} \quad F_{s2} = k_2(x_{s2} - x_{s1}) \tag{12.27}$$

The $x_{si}$ in eq. (12.27) are absolute $x$-axis measurements; $x_{ser} = x_{s2}$ is the movement of node 2 where the load $P$ is applied. Each spring in a series array carries the same force, $F_{s1} = F_{s2} = P$. We can then write a force balance equation at points 1 and 2 to easily calculate the total deflection of the endpoint at which $P$ is applied:

$$x_{ser} = \left(\frac{1}{k_1} + \frac{1}{k_2}\right)P = \left(\frac{k_1 + k_2}{k_1 k_2}\right)P \tag{12.28}$$

The *effective stiffness* of a pair of spring in series is then

$$\frac{1}{k_{ser}} = \frac{1}{k_1} + \frac{1}{k_2} = \frac{k_1 + k_2}{k_1 k_2} \tag{12.29}$$

We also note that the net extension of spring 2, $\delta_{s2} = x_{ser} - x_{s1}$, can be calculated from eqs. (12.27) and (12.28) as

$$\delta_{s2} = \left(\frac{k_1}{k_1 + k_2}\right)x_{s2} \tag{12.30}$$

Equation (12.30) is interesting because it shows that a series combination of springs serves as a *displacement divider* because the net extension $\delta_{s2}$ of spring 2 is a fraction of the total

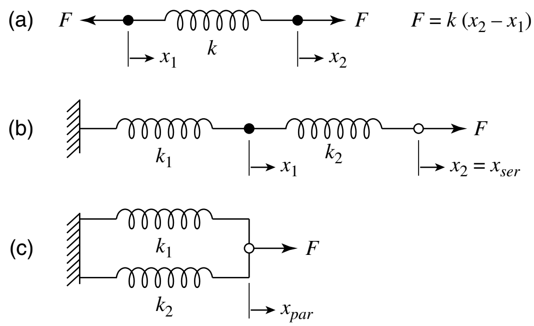



**Figure 12.1** Springs as one building block for mechanical systems or circuits: (a) a spring's constitutive law; (b) two springs in series; and (c) two springs in parallel.

extension $x_{s2}$, and that fraction depends (here) on the ratio of the stiffness of spring 1 to the sum of the stiffnesses. Thus, if spring 1 is very stiff, it will be spring 2 that extends the most. If spring 2 is relatively stiff, then it will extend less and spring 1 will extend more. This is in accord with our intuition: Imagine a rubber rod in series with a steel rod! All of the foregoing results are readily extended to $n$ springs in series.

A similar set of calculations for the common displacement and the forces in two springs connected in parallel produces similar kinds of results. The common extension of the two parallel springs is

$$x_{par} = \frac{P}{k_1 + k_2} \tag{12.31}$$

The spring stiffnesses add linearly when they're in parallel, so that their effective stiffness is

$$k_{par} = k_1 + k_2 \tag{12.32}$$

However, and unlike its series counterpart, the forces carried by each spring in a parallel array are not the same:

$$F_{p1} = \frac{k_2 P}{k_1 + k_2} \quad \text{and} \quad F_{p2} = \frac{k_1 P}{k_1 + k_2} \tag{12.33}$$

We note from eq. (12.33) that a pair of springs in parallel serves as a *force divider* because the force in each spring is a fraction of the total force that is determined by the ratio of its stiffness to the system's effective stiffness.

The elementary electrical circuits shown in Figure 12.2 can also be viewed through the series–parallel prism. Our basic constitutive law is *Ohm's law*: A current $I$ flowing through a resistor $R$ produces a voltage drop across that resistor (see Figure 12.2(a)):

$$I = \frac{1}{R}(V_a - V_b) = \frac{1}{R}V \tag{12.34}$$

(The voltage is measured in volts (V), the current in amperes (A), and the resistance in ohms (Ω).) Figure 12.2(b) shows a voltage $V_0$ applied across two resistors in series. Just

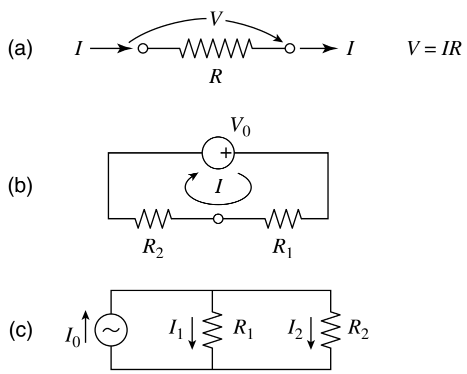



**Figure 12.2** Resistors as one building block for electrical circuits or systems: (a) a resistor's constitutive law; (b) two resistors in series; and (c) two resistors in parallel.

as two springs in series carry the same force, the two resistors in series have the same current flowing through each. The voltage drop across the two resistors in series $V_{ser}$ must equal the input voltage $V_0$ across them, that is,

$$V_{ser} = V_0 = V_1 + V_2 = IR_1 + IR_2 = I(R_1 + R_2) \tag{12.35}$$

Thus, the *effective resistance* of two resistors in series is simply the sum of the individual resistances:

$$R_{ser} = R_1 + R_2 \tag{12.36}$$

The voltage drop across each of the resistors can be written as a fraction of the input voltage:

$$V_1 = \frac{R_1}{R_1 + R_2} V_0 \quad \text{and} \quad V_2 = \frac{R_2}{R_1 + R_2} V_0 \tag{12.37}$$

Thus, the two resistors in series serve as a *voltage divider*, reminiscent of the two springs in series acting as displacements dividers.

We can also make similar observations about the two resistors connected in parallel, as shown in Figure 12.2(c). Here we view the input as a current source, especially since the parallel nature of the circuit requires that the voltage drop across each resistor is the same as the voltage drop across that current source. This is akin to the common extension shared by two springs in parallel. What is of interest here is that the current $I_0$ provided by the current source divides into two components $I_0 = I_1 + I_2$ that share a common voltage drop $V_{par}$:

$$V_{par} = I_1 R_1 = I_2 R_2 \tag{12.38}$$

Expressed in terms of the current input, the voltage drop across this parallel circuit is

$$V_{par} = \left(\frac{R_1 R_2}{R_1 + R_2}\right) I_0 = \left(\frac{1}{R_1} + \frac{1}{R_2}\right)^{-1} I_0 \tag{12.39}$$

Thus, the effective resistance of this parallel circuit is

$$\frac{1}{R_{par}} = \frac{1}{R_1} + \frac{1}{R_2} = \frac{R_1 + R_2}{R_1 R_2} \tag{12.40}$$

Finally, this parallel connection of two resistors acts as a *current divider*, as we see when we write the currents through each resistor as a fraction of the current source input:

$$I_1 = \frac{R_2}{R_1 + R_2} I_0 \quad \text{and} \quad I_2 = \frac{R_1}{R_1 + R_2} I_0 \tag{12.41}$$

Thus, if we compare eq. (12.41) to eq. (12.33), we see that the current divider in a parallel circuit acts very much like the force divider on the parallel *mechanical circuit* of two mechanical springs acting in parallel. In that vein, such dividers also exist for other linear elements in mechanical systems (e.g., dampers) and electrical circuits (e.g., capacitors and inductors).

We should not mistake the foregoing discussion as anything even remotely close to complete analyses of mechanical or electrical circuits or systems. We've shown these calculations only to point out the value of looking for series and parallel connections, whether in circuits or in very complicated systems—which can often be modeled in terms of sophisticated, lumped elements. Such series–parallel assessments provide an elegant and intuitive way to track the flow of variables *through* system elements (e.g., force in mechanical systems and current in electrical circuits) and changes of variables *across* circuit elements (e.g., displacements or movements in mechanical systems and voltage in electrical circuits). We see this clearly in the notion of *dividers* that apportion various inputs corresponding to element properties, a notion that is particularly interesting when sorting out functions.

### 12.2.7 Mechanical–Electrical Analogies

We pointed out the strong similarity between the elementary mechanical and electrical circuits we've just analyzed, which suggests there may be an analogy between the two. There is, however, also one notable difference: The mechanical elements are springs, which store energy, while the electrical elements are resistors, which actually dissipate energy (as waste heat). Does this reduce the likelihood of an analogy or diminish its utility even if there is one? In fact, there are other mechanical–electrical analogies, and they're not equally useful.

We also pointed out that the series–parallel approach applied to other elements, such as the capacitor, which has the constitutive relation

$$Q = \frac{1}{C}(V_a - V_b) \quad \text{or} \quad I \triangleq \frac{dQ}{dt} \triangleq \dot{Q} = \frac{1}{C}(\dot{V}_a - \dot{V}_b) \tag{12.42}$$

(The capacitance $C$ is measured in Farads (F).) We see that equations in (12.42) strongly resemble their spring counterparts in (12.27), and this observation lies at the heart of the best mechanical–electrical analogy. We can liken current to force as variables that *flow through* elements, and voltage drops to the difference in endpoint displacements as variables that *change across* an element. And both capacitors and springs store energy.

Further, when we write equilibrium equations, or balance forces, we are actually conserving momentum. The corresponding electrical rule is *Kirchhoff's current law*, which conserves current by balancing all of the currents coming into and going out of a junction or node. Similarly, when we ensure that all of the individual spring endpoint displacements properly add up over a chain of springs, we are ensuring spatial consistency or compatibility. The electrical analogy is *Kirchhoff's voltage law*, which says that the sum of voltage drops around a closed circuit loop must be zero. Similarly, just as there are energy storage elements (mechanical springs, electrical capacitors and inductors), there are also energy dissipaters (mechanical dampers, electrical resistors).

A complete coverage of all of the nuances (and potential) of the mechanical–electrical analogy is well beyond our scope. However, such analogical awareness is another good habit of thought that experienced designers often exploit.

## 12.3 MODELING A BATTERY-POWERED PAYLOAD CART

We now illustrate how we might do some basic mathematical modeling for a fairly common design problem. In so doing, we will focus on how we formulate or set up mathematical models, rather than on solving them. Also, as we model this cart-payload-ramp design problem, we will post signs [***in brackets***] to indicate which modeling principles (as represented by the questions in Section 12.1.1) we are applying.

Suppose we are asked to design a cart to maximize the height $h$ to which a payload of weight $W_p$ can be moved up a ramp at an angle $\theta$ with the horizontal (Figure 12.3). [***Why***] This is a challenge on several fronts: How do we design a cart to contain and support the payload? Can we model a cart without having done its detailed design? How do we provide power to that cart, that is, how do select the appropriate battery and motor? (Note that we have implied a design solution here by assuming we're using a battery and a motor: We could perform the function of providing power with a spring.) [***Assume***] We'll subsume the details the cart's structural form into a parameter $W_c$ that reflects the total weight of the cart and its battery and motor. [***Assume***] The basic purpose we want our model to serve is to define how much power will be needed to move the cart up the ramp. In other words, we want to identify how much electrical energy we need to transform into mechanical energy to move the cart up the ramp. [***Use***]

To develop a useful model, we must first decide what system(s) we want to model, and in how much detail. There are at least three ways to configure a system model [***How***]: The cart, battery, and motor are considered a single system within the boundary shown in Figure 12.4(a); the overall system model is decomposed into two *subsystems* of cart and of battery and motor, as shown in Figure 12.4(b); the overall system model is further decomposed into three subsystems of cart, battery, and motor, as shown in Figure 12.4(c). The first two models differ only if we want to distinguish between batteries that are chargeable and those that are not. Since that's a relatively unimportant difference at this stage, we will go with the two-subsystem model, wherein electrical energy is the input to the first stage, either as a new battery or as a line cord to recharge the chargeable batteries. [***Given***] We will formulate the overall system shown in Figure 12.4(a), and then we'll describe the virtues and the shortcomings of that model. After that, we'll provide some high-level considerations for choosing batteries and motors.

### 12.3.1 Modeling the Mechanics of Moving a Payload Cart up a Ramp

From the overall systems and mechanics points of view, the system inside the boundary of Figure 12.4(a) can be modeled rather simply: Consider a "particle" or lumped mass of the

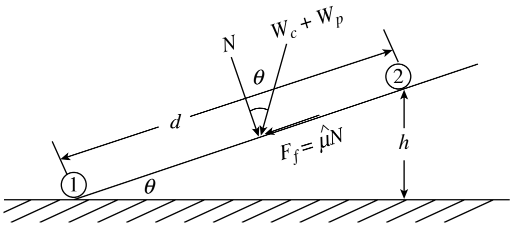



**Figure 12.3** A sketch of the geometry and principal forces involved in modeling how a battery-powered cart moves a payload up a ramp.

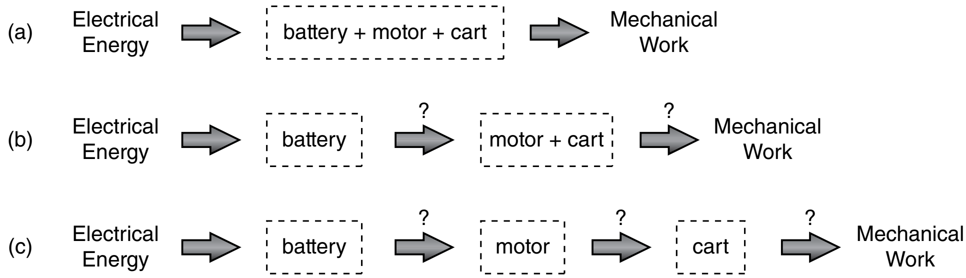



**Figure 12.4** Different models of a battery energizing a motor that is in turn powering a cart require different system boundaries (shown as dashed lines): (a) a single (integrated) system; (b) a two subsystem model of (battery + motor) and the cart; and (c) a three subsystem model.

cart $(W_c/g)$ and payload $\left(W_p/g\right)$ at rest at the foot of the ramp (position 1 in Figure 12.3). Electrical energy is the input to the system (assuming an electrical cord to recharge a chargeable battery) that is transformed into mechanical energy, which is then used to move the cart up the ramp. Some of that energy will be lost due to friction, and we'll address that shortly. [***Why***] So our starting point is the conservation principle, expressed in the work–energy relationship of eq. (12.26) and repeated here [***How***]:

$$\left[\frac{1}{2}mv_2^2 - \frac{1}{2}mv_1^2\right] + V_2 - V_1 = \int_1^2 \vec{F}\cdot d\,\vec{s} \qquad (12.26)$$

The force $\vec{F}$ in the work integral in eq. (12.26) is solely dissipative because any conservative forces would be represented in appropriate potential energy terms. In the present case, both of the kinetic energy terms vanish since the cart starts and ends at rest [***Assume***]. Measured from a datum on the surface at the foot of the ramp, the potential energy at position 2 is

$$(PE)_2 = \left(W_c + W_p\right)h = \left(W_c + W_p\right)d\sin\theta \qquad (12.43)$$

The gravitational potential energy at position 1 is zero since the cart is at the datum. However, this is the logical and convenient place to introduce the energy input, that is, the energy $E_b$ stored in the battery. [***Given***] Were this a spring-loaded cart, we'd clearly write $V_1$ as the energy stored in an elastic spring of some (presumably given) stiffness $k$. Here the energy is stored in a battery:

$$V_1 = E_b \qquad (12.44)$$

We should note here that we must be careful when assigning a value to $E_b$ simply because batteries and the motors they drive are not 100% efficient. While motors often are assigned efficiency ratios $\eta$ (as we discuss in Section 12.3.3), it may be easier at this level of detail to reduce the value of $E_b$ by some fraction to account for motor losses.

As we model the work done to move the cart and its payload up the ramp, we should keep in mind that friction inevitably works against us, the second law of thermodynamics being what it is [***Given***]. For our cart, work is done against both internal friction (e.g., in the motor's gears) and external friction (e.g., between the ramp surface and both the driver wheels as rolling resistance and the passive driven wheels as sliding friction), until the cart comes to a stop. At some point the cart will stop because we have used some of our initial energy to increase the potential energy of the cart and its payload by raising their height, and because of the aforementioned friction loss. The overall system model we're using here allows us to "disguise" the details of the different kinds friction by assuming that all of the work done against friction can be lumped together and accounted for in a simple product of the normal force of the cart (produced by the cart and payload) multiplied by a coefficient of sliding friction $\hat{\mu}$ [***Assume***]:

$$\int_1^2 \vec{F} \cdot d\vec{s} = -\hat{\mu}\left(\left(W_c + W_p\right)\cos\theta\right)d \tag{12.45}$$

Then, if we substitute eqs. (12.43)–(12.45) into eq. (12.26), we find the overall work–energy relationship can be put in the following form [***Predict***]:

$$E_b = \left(W_c + W_p\right)d(\sin\theta + \hat{\mu}\cos\theta) \tag{12.46}$$

Equation (12.46) is interesting and seems consistent with many of our intuitions. First, the physical dimensions of both sides of eq. (12.46) are those of work, as appropriate. Second, we see that for a given amount of battery energy, there is a clear trade-off between the total weight $\left(W_c + W_p\right)$ and the distance $d$ that the cart can move up the ramp. We make that trade-off more apparent by recasting eq. (12.46) to calculate the actual distance the cart moves:

$$d = \frac{E_b}{\left(W_c + W_p\right)(\sin\theta + \hat{\mu}\cos\theta)} \tag{12.47}$$

We can equivalently calculate the height $h$ to which the loaded cart is raised:

$$h = d\sin\theta = \frac{E_b}{\left(W_c + W_p\right)(1 + \hat{\mu}\cot\theta)} \tag{12.48}$$

Both eqs. (12.47) and (12.48) are dimensionally homogeneous. Further, there are two limiting cases that produce results consistent with our intuitions. If the system were ideal and frictionless (i.e., $\hat{\mu} = 0$), the cart would move up the ramp up to a limiting height for which all of the battery energy has been converted into gravitational potential energy. Similarly, if there was no ramp (i.e., $\theta = 0$), eq. (12.48) tells us that $h = 0$, while eq. (12.47) tells us that on a level surface, the loaded cart would travel a distance

$$d_{\theta=0} = \frac{E_b}{\hat{\mu}\left(W_c + W_p\right)} \tag{12.49}$$

We should keep in mind that while eqs. (12.47)–(12.49) have the virtues of dimensional consistency and reasonable limiting behavior, they also embody a lot of assumptions about the role of friction in this problem. First of all, we have no idea of what the friction coefficient $\hat{\mu}$ really means or what its value might be. In fact, we would likely want to do both deeper analyses and some significant physical testing to see whether the basic assumptions of eq. (12.45) can be validated [***Validate***] or how they should be modified if not [***Improve***].

Further, the overall systems boundary and the simple assumption about friction mask a much more complicated problem. There are a lot of powered toy cars that we can play with—and likely have. So chances are we've seen that the driver wheels of such a toy car often spin at start-up, while the car doesn't move for some short period of time. Then the car seems to "get a grip" and take off, but often the driven wheels don't seem to turn as much as they seem to slide. This is actually a complicated mechanics problem. To start from rest, the powered driver wheels must provide enough torque to overcome static friction, which is larger than rolling resistance or sliding friction. But if too much torque is provided initially, then the wheels spin without "grabbing" the surface, which is why we see what we've just described. That means the starting torque must fall off to get the cart moving, and when the car is actually moving, we need a more complex model to describe its behavior. That in turn means that we need a more detailed mechanics model: It also means that we have to delve more deeply into how motors actually work, that is, how the torque they produce varies with the speed at which their armatures rotate. To get into the details of such modeling, we would have to work within the three subsystem boundaries defined in Figure 12.4(c) [***Improve***]. We won't, since that's beyond our scope, but we can't resist observing that the kind of behavior we've been describing is often seen when drivers, intentionally or not, gun their engines while starting their cars.

Before we turn to choosing batteries and motors, it is interesting to cast this design problem in the context of a design challenge or contest. Because of the weight–distance trade-offs we mentioned earlier, if we wanted to compare different cart designs, one against another, we might phrase a challenge in terms of maximizing a *figure of merit* (FoM) such as [***Predict***]:

$$\text{FoM} = d\left(1 + W_p/W_c\right) \tag{12.50}$$

Such a figure of merit emphasizes how the cart designers would have to trade off the distance they hope to achieve against the payload they can carry, with that payload normalized against the cart weight—which includes weights of both a battery and a motor. So there might be further refinements if, as we might expect, the available battery energy varies with battery size, and thus battery weight.

We also note that our cart design problem was stated independent of time. But suppose we wanted to get to a certain height above the datum in a specified time, which would serve as a design constraint. Or imagine that we wanted to minimize the time it would take to climb the ramp, in which case minimizing transit time would be an objective. Both of these challenges could be reflected in a revised figure of merit, namely, if we modified eq. (12.50) to read [***Predict***]:

$$\text{FoM}' = d\left(1 + W_p/W_c\right)/t \tag{12.51}$$

In this case, we would not only need to find enough energy to move the cart, but we would also have to examine how that energy is delivered over some interval of time, that is, we need to examine its power. Thus, in searching for a battery and a motor, we need to evaluate both the time variation of a battery's energy and the operating characteristics of the motor.

### 12.3.2 Selecting a Battery and Battery Operating Characteristics

We start with batteries. How much energy is stored in a battery? [***Why***] This question is not easily answered, especially when we're buying them off a shelf. When batteries discharge their stored energy, they convert chemical energy to electrical energy as inputs to circuits. A battery can be characterized in terms of:

- *voltage* $V$ (V);
- *energy density* $E_b$ [typically measured in terms of either watt-hours/kilogram (W h/kg) or watt-hours/liter (W h/l)];
- *capacity* $Q$ [typically expressed in ampere-hours (A h) or milliampere-hours (mA h)]; and
- the *shape of the battery's discharge curve*.

Depending on the particular battery, very little of this data is actually given: We most often buy batteries by their nominal voltage (e.g., 1.5 V), which is often expressed in terms of a rated physical size (e.g., AA). Battery capacity is a most important characteristic, especially when we're *sizing* a battery to provide a requisite power over a desired time span. In our cart-payload-ramp problem, we've not made time a particularly important variable or parameter, so we just want to know how much energy we can tap, rather than how fast or at what rate the energy is available. We can estimate that available energy from a battery's *discharge curve*, which we also call its *operating characteristics*.

We show two sets of discharge curves in Figure 12.5, along with an idealized, archetypal curve that we will use to highlight our estimates of battery capacity. Figures 12.5(a) and (b) show *measured* plots of voltage against time (they are readily found on the World Wide Web), and they show different curves for different current draws: The voltage drop takes longer when the current draw is lower. We also intuit that these curves reflect somehow how much energy is available. Since all of the data in these discharge curves is empirical, we need some easy estimators to get an idea of just how much energy is available, hence we use the idealized discharge curve in Figure 12.5(c) [***Assume***]. For that discharge curve, we identify the capacity $Q$ as the total charge in the battery [***How***]. That capacity can be estimated by multiplying the draining current $I_s$ by the time it takes to reach a precipitous drop-off in the voltage curve $t_d$:

$$Q \cong I_s t_d \tag{12.52}$$

The physical dimensions of $Q$ are charge, typically measured in coulombs (C), and those of the current $I_s$ are charge per unit time, with normal units of amperes (A = C/s). However, because of the ranges of the numbers involved, battery capacity is typically expressed in units of ampere-hours (A h = A s/3600) or milliampere-hours (mA h). So, depending on both the resistance through which the battery drains and its discharge current, Figure 12.5(a) suggests (roughly) that $Q_{AA} = (1.4)(1) = 1.4, (3.6)(0.5) = 1.8$ or $(8.2)\ (0.25) = 2.05$ A h [***Find***]. The capacity of the LR44 is noticeably lower, as can be estimated from Figure 12.5(b): $Q_{LR44} = (675)\left(181.8 \times 10^{-6}\right) = 0.122$ A h [***Find***]. An LR44 battery has less capacity than an AA, draws much less current, and is physically much smaller than a AA. That is why LR44s are used in devices where they are exchanged infrequently, for example, watches and calculators.

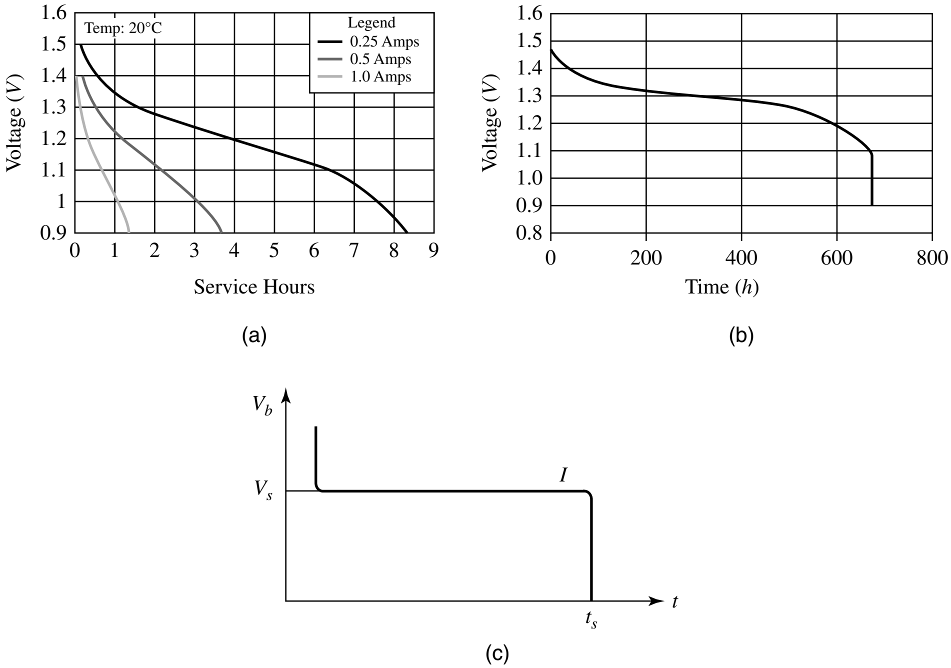



**Figure 12.5** Battery discharge curves or battery operating characteristics: (a) AAs for different currents; (b) LR44 at 181.8 μA; and (c) an idealized, archetypal discharge curve with nominal service voltage $V_s$, service time $t_s$, and draining current $I_d$.

The energy stored in a battery can then be estimated as [***Predict***]:

$$E_b \cong V_s Q = V_s I_s t_d \tag{12.53}$$

In eq. (12.53) we have introduced the average service voltage $V_s$, which we usually estimate by looking for the plateau in a battery's discharge curve. Further, it is quite clear from the idealized battery discharge curve (Figure 12.5(c)) that eq. (12.53) is an estimate of the area under the discharge curve. $E_b$ has the physical dimensions of voltage × current × time or voltage × charge, which are the appropriate dimensions for electrical work. For the case of the LR44 battery, whose operating capacity is shown in Figure 12.5(b), that plateau is fairly evident, $V_s \cong 1.3$ V, and so $E_{LR44} \cong (1.3)(122) \cong 159$ W h [***Find***].

We should use the calculated value of $E_b$ with some caution, principally because we are not accounting for any losses in the battery itself or in the ways we connect it to whatever device we use to do the mechanical work we want done [***Why***]. The simplest way to model this aspect is to recognize that a battery is an *emf*: A device that converts chemical, mechanical or some other from of energy into electrical energy. It has a (voltage) value $V_b$. The emf is connected in series to the battery's own internal resistance $R_b$ and to another resistor $R$ that represents a load (see Figure 12.6) [***How***]. That resistor could represent a simple way to dissipate energy through a resistance load, or it could be a simple stand-in for

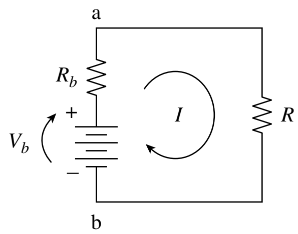



**Figure 12.6** Circuit of a battery with emf $V_b$ and internal resistance $R_b$ connected to a load resistor $R$, which can also serve as an elementary model for the load imposed by a motor.

a motor with the power, otherwise dissipated, being fed into some transmission device to do mechanical work [***Assume***]. In our case, we would regard that power "loss" as providing the energy needed to do the work of moving the loaded cart up the ramp.

The circuit in Figure 12.6 is relatively simple. It has a single loop current $I$, and we can analyze it by balancing the voltage drop across the two points $a$ and $b$, which yields the following instance of Kirchhoff's voltage law [***Predict***]:

$$V_b - IR_b = IR \tag{12.54}$$

Since $R_b$ represents the battery's internal voltage, we may regard the term $(V_b - IR_b)$ as the net voltage available at the battery's terminals (think car batteries!). The power output across the resistance $R$ is [***Predict***]

$$P_{out} = I^2 R = \frac{V_b^2 R}{(R + R_b)^2} \tag{12.55}$$

We obtain the second form of eq. (12.55) by eliminating the current $I$, which we found from the Kirchhoff statement (12.54). The physical dimensions of power are those of work/time, and the corresponding metric units are watts or joules/seconds (W = J/s).

Remembering that we regard this as useful power output in our elementary model of a motor as the resistor, we would like to develop the motor's maximum power. Thus, using classical calculus, we evaluate

$$\frac{dP_{out}}{dR} = 0 \tag{12.56}$$

We find the maximum power is developed when the motor's resistance equals that of the battery (i.e., $R = R_b$), and that maximum power is

$$P_{max} = P_{out}\big|_{R=R_b} = \frac{V_b^2}{4R_b} \tag{12.57}$$

There will inevitably be losses in a real motor, and in any attached transmission components (e.g., gears). So, it would be extraordinarily optimistic to simply accept the calculated values of $E_b$ (eq. (12.53)) and $P_{max}$ (eq. (12.57)) without considering how both the battery and the motor are actually situated and used in the overall cart design.

### 12.3.3 Selecting a Motor and Motor Operating Characteristics

We finished our description of battery behavior by modeling a motor as a resistor in series with a battery. That's all well and good as far as it goes, but it is also useful to think a little about how a simple motor operates, particularly in terms of its output. A motor can be characterized in terms of:

- *angular speed* $\omega$ [typically measured in units of revolutions per minute (rpm)];
- *torque* $T$ [with physical dimensions of F·L and measured in the metric system in units of (N m) and in the British system in ($lb_f$ ft)];
- *gear ratio* $n$ [dimensionless ratio of the number of teeth on the driving gear to the number of teeth on the driven (output) gear];
- *power* $P$ [with physical dimensions of F·L/T and measured in the metric system in units of (W or N m/s) and in the British system in (hp or $lb_f$ ft/sec)];
- *motor efficiency* $\eta$ [dimensionless ratio $P_{out}/P_{in}$]; and
- the *shape of the motor's characteristic operating curve*.

As we noted just above, a major descriptor of a motor is the shape of its characteristic operating curve, which is a plot of the motor torque against its rotational speed. For DC motors generally, the relationship of torque to rotational speed is given by:

$$T(\omega) = T_s\left(1 - \frac{\omega}{\omega_0}\right) \tag{12.58}$$

Equation (12.58) is plotted in Figure 12.7, with the aid of which we can see why $T_s$ is called the *stall torque*: it is the torque at which the motor stalls or stops turning (i.e., $\omega = 0$); and $\omega_0$ is called the *no-load speed* because it is the speed at which there is no load on the motor (i.e., $T = 0$). Further, the power produced by such a DC motor can be calculated as:

$$P_{out}(\omega) = \omega T(\omega) = T_s\left(1 - \frac{\omega}{\omega_0}\right)\omega \tag{12.59}$$

From eq. (12.59) we can easily show that the motor produces a maximum power at one-half of the no-load speed with the corresponding torque being one-half of the stall

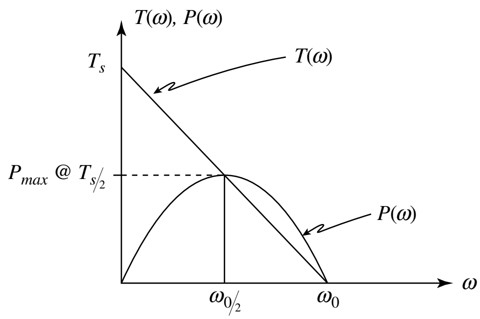



**Figure 12.7** The torque vs. speed operating characteristic curve for a DC motor.

**TABLE 12.3 Operator characteristics of typical geared DC motors, arranged in order of their decreasing gear ratios**

| Type | Gear Ratio ($n$) | Torque ($T$; gf cm) | Torque ($T$; N m) | Angular Speed ($\omega$; rpm) |
|---|---|---|---|---|
| A | 344.2 | 2276 | 0.2230 | 38 |
| B | 114.7 | 809 | 0.0793 | 115 |
| C | 38.2 | 278 | 0.0272 | 345 |
| D | 12.7 | 94 | 0.0921 | 1039 |

torque, that is,

$$\frac{dP_{out}(\omega)}{d\omega} = 0 \Rightarrow P_{max}(\omega) = \frac{\omega_0 T_s}{4} \tag{12.60}$$

Practically speaking, experts employ some rules of thumb aimed at maximizing motor efficiency when they are choosing motors to provide a specific torque and a specific operating speed. Those heuristics are that the motor:

- should run at 70–90% of the no-load speed ($\omega_0$); and
- should provide a torque of 10–30% of the stall torque ($T_s$).

Generally, when we choose a motor to provide a specific torque while running at a specific speed, we also need to account for the gearing that comes along with the motor because the *gear ratio* is a property of that motor. In Table 12.3 we show data for a small, geared DC motor, drawn from a manufacturer's catalogue. We've extended the manufacturer's data to show two sets of torque values. The first set shows $T$ values in nonstandard units of gf cm, a practice that is not uncommon in many such catalogues. Why would someone use such nonstandard units? Perhaps the answer follows from the second set of $T$ data, expressed in the more customary N m, and for which the numbers are small: When doing quick hand calculations, which we often do in this kind of modeling, it's easier to deal with the numbers that arise when $T$ is measured in those nonstandard units. Note, too, that the data as given (and extended) do not explicitly identify the power outputs of the four motors listed.

The power produced by a motor is simply the product of the torque and the angular speed. In standard metric units the power is

$$P_{out} = P(\mathrm{W}) = \frac{\omega T}{9.554} \tag{12.61}$$

In conventional American units the power is

$$P_{out} = P(\mathrm{hp}) = \frac{\omega T}{5252} \tag{12.62}$$

We have calculated the power output of each of these motors and show the data in Table 12.4. Clearly, the four motors listed all produce *almost* the same (small) amount of power, at an average of $P_{out/ave} = 0.96\ \mathrm{W} = 0.0018\ \mathrm{hp}$. However, we can see by reading up

**TABLE 12.4 Power outputs of typical DC motors, arranged in order of their decreasing gear ratios**

| Type | Gear Ratio ($n$) | Torque ($T$; N m) | Angular Speed ($\omega$; rpm) | Power (W) | Power (hp) |
|---|---|---|---|---|---|
| A | 344.2 | 0.2230 | 38 | 0.88 | 0.0016 |
| B | 114.7 | 0.0793 | 115 | 0.95 | 0.0017 |
| C | 38.2 | 0.0272 | 345 | 0.98 | 0.0018 |
| D | 12.7 | 0.0921 | 1039 | 1.05 | 0.0019 |

the table (from the bottom row) that the power drops off as the gear ratio increases. This is common and perhaps ought to be another motor heuristic: an increase in the gear ratio will produce a decrease in the power delivered. Thus, while we can provide more torque by using a higher gear ratio, it comes at the cost of both a lower speed and less power (i.e., fewer watts or lower horsepower numbers). These results seem intuitively right.

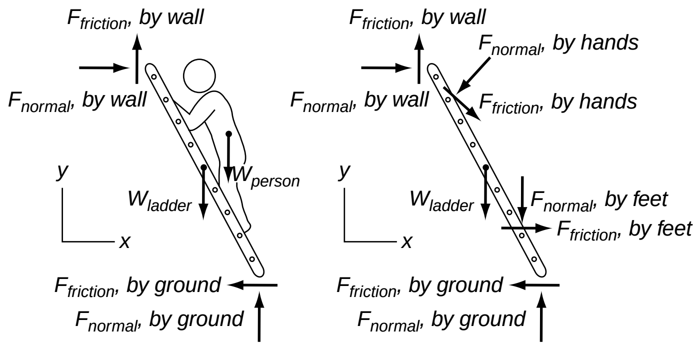
It also turns out that many of the small motors characterized as those in Table 12.3 are designed to work with batteries of a particular voltage. Thus, all of the motors in Table 12.3 run on 3 V batteries. So the battery–motor combination is fixed, once we choose such a motor.

Finally, although this would not likely come into play for the kinds of small toy motorized carts we have been talking about, it turns out that one major issue in motor selection is that they generate a fair amount of heat. The amount of heat generated will be a consequence of the current $I$ through the motor's resistance $R$, as specified by eq. (12.55). Again, for small-scale models and toys, this should not be an issue since we are talking about very small amounts of power.

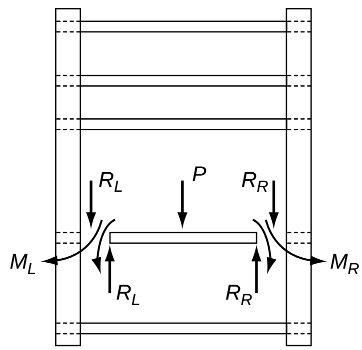
We now know the principles of sizing a battery and choosing a motor, so we can design a cart—at least to the extent of balancing the payload-to-cart weight ratio against the distance $d$ moved up the ramp, as a function of the chosen $E_b$, and as a function of $P_{out}$ if time is important in our design. We may also view the battery as a constraint, for example, when only certain batteries are available, or perhaps because of cost considerations. And, of course, we have to remember that the physical size of the battery and the motor will also enter into our design considerations because its weight is part of the overall cart weight.

## 12.4 DESIGN MODELING OF A LADDER RUNG

We now model and design the step or rung of a ladder to show how modeling is needed to do design properly. In order to design a ladder's step we need a model that predicts its behavior [***Why***], which means that we want to be able to understand how the step's attributes (e.g., size, shape, material, connections to the ladder's frame) affect its ability to support given loads [***Find***]. We are told that the ladder must support a person of specified weight $W_p$ carrying a specified weight $W_w$ [***Given***]. We will model the behavior of the step (and the ladder) using a standard model of linear elastic beam behavior (described in Section 12.4.1) and standard models of the elastic behavior of materials [***Assume***]. We will develop and apply the model of the step using basic principles of mechanics [***How***] and




(a)




(b)

**Figure 12.8** *Free-body diagrams* (FBDs) of various aspects of a person standing on a ladder: (a) side-elevation FBDs of person and ladder taken as a system; (b) a front-elevation FBD of the ladder showing vectors of the force $F_{normal} = P = W_p + W_w$ due to the person standing on a rung, as well as the vertical forces ($R_L$, $R_R$) and *moments* ($M_L$, $M_R$) by which the ladder's side rails support the rung. Courtesy of S. D. Sheppard and B. H. Tongue. Reprinted by permission of John Wiley & Sons, Inc.

show just how the total weight that it can support depends on its geometric and material properties [***Predict***]. Note how we are limiting our modeling effort: At this point we are *not* analyzing the size, shape, or materials of the side frames, any cross bracing, or any footpads on a rung. We are also precluding any linkages among the various ladder supports. One or more mathematical models would be needed to develop these parts of the ladder, but we will just model an individual rung or step.

Let us now apply some basic mechanics principles. In Figure 12.8 we show three sketches of a person on a ladder, the first of which is a *free-body diagram* (FBD) of the person and the ladder taken as a system. The second sketch shows an FBD of the entire ladder. The third drawing shows elevations of FBDs of the rung, on which are shown vectors representing (a) the force exerted by the load-carrying person and (b) the vertical forces and *moments* provided by the ladder's frame to support the step.

### 12.4.1 Modeling a Ladder Rung as an Elementary Beam

Imagine the two scenarios shown in Figure 12.9, both of which show a vertical load $P$ supported by a transverse (i.e., normal or, here, horizontal) element. Figure 12.9(a) shows a *cable* or rope, along with FBDs of two sections of the cable. We see that the load seems to make the cable kink, and that a vertical load can be supported by a tensile force $T$ in the cable or rope. In Figure 12.9(b) we show a *beam*, along with two FBDs of the beam divided into two sections. In the first FBD, we see that the external vertical force in each section is supported by reactions $R_A$ and $R_B$ at each beam support and an internally developed *shear force*, $V$. However, there is nothing to prevent either section from rotating or spinning because each has an unbalanced couple or *moment* in the configuration shown. In the second beam FBD we have included *bending moments* $M$ that are internally developed couples (or moments) that maintain moment equilibrium and thus prevent each section from spinning out of control. These bending moments are developed by in-plane stresses along the axis of the beam, so that it is a set of *horizontal* stresses that support a *vertical* load in a beam!

For our purposes, the important aspects of elementary beam theory are that beams behave like linear springs and the stiffness of a beam depends on several of the beam's parameters. For a simple beam, a load $P$ applied at the mid-way point of a beam of length $L$ produces a deflection $\delta$ of the beam under that point:

$$\delta = \frac{PL^3}{C_\delta EI} \tag{12.63a}$$

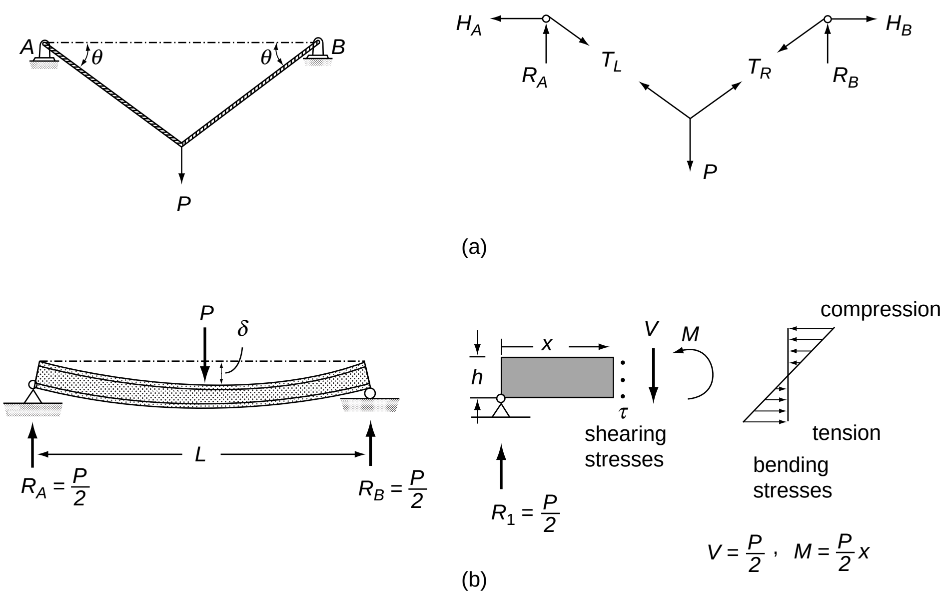



**Figure 12.9** Supporting a vertical force with a transverse (horizontal) structure: (a) *cable* and FBDs of two sections of the cable; and (b) a *beam* and a FBD of one section of the beam that also shows how a moment (couple) is developed by axially directed normal stresses on the beam's cross-sectional area.

Here $C_\delta$ is a number that changes with the beam's boundary conditions. (We used dimensional analysis to derive this result as eq. (12.13), and showed the dimensions of the variables involved in Table 12.2.) We can also rewrite eq. (12.63a) as an analog of the classical spring formula, $F = kx$, that is,

$$P = \left(\frac{C_\delta EI}{L^3}\right)\delta \tag{12.63b}$$

The other physical quantity of great interest in beam theory is the bending stress along the axis of the beam. As may be visualized from Figure 12.9(b), it is the bending stress that creates the bending moment and its consequent shear force that enables a long thin beam to support a load that acts in a direction normal to the (long) axis of that beam. The maximum stress in a loaded beam is

$$\sigma = \frac{PLh}{2C_\sigma I} \tag{12.64}$$

Here $h$ is the height of the beam's cross-section (see, again, Figure 12.9) and $C_\sigma$ is a number that changes with the beam's boundary conditions. Stress has the same physical dimensions as pressure, that is, $[\sigma] = \mathrm{F/L^2}$.

The classical spring formula has only one constant or design variable that can be chosen or manipulated, $k$, which thus limits its design freedom. For the beam that has to span a given length $L$, there are three variables that can be varied: $E$, $I$, and $h$. (To the extent we can choose how the beam is supported at its ends, we can also choose between appropriate pairs of constants, $C_\delta$ and $C_\sigma$). The increased number of variables means that we can design to achieve objectives or constraints expressed in terms of the beam's deflection (eq. (12.63a)) and its maximum stress (eq. (12.64)). Thus, we shall soon talk about *designing the beam for stiffness*, when the deflection is our focus, or *designing the beam for strength*, when the maximum stress is our focus.

We see in Figure 12.10 that there are different rung supports (or connections) to stipulate or model. The two limiting cases that are of most relevance are pictured in Figure 12.11: *Simple* (or *pinned* or *hinged*) supports that provide a vertical reaction force and prevent any vertical deflection but leave the ends of the beam free to rotate; and *fixed* (or *rigid* or *clamped*) supports that provide both vertical reaction forces that prevent vertical deflections and moments that force the slope of the deflection of the beam to vanish (i.e., the moments prevent any rotation at that support). These two limiting cases, informed by our actual experience, suggest that we will have to make another modeling-design assumption when we design the rung.

With eqs. (12.63) and (12.64) in hand, and a tentative decision taken about the kinds of beam supports we will consider, we have specified the equations we will use, the calculations we can make and the types of answers we might expect [***Predict***]. Thus, we have established a principled model that we can now use for the preliminary design of a ladder rung.

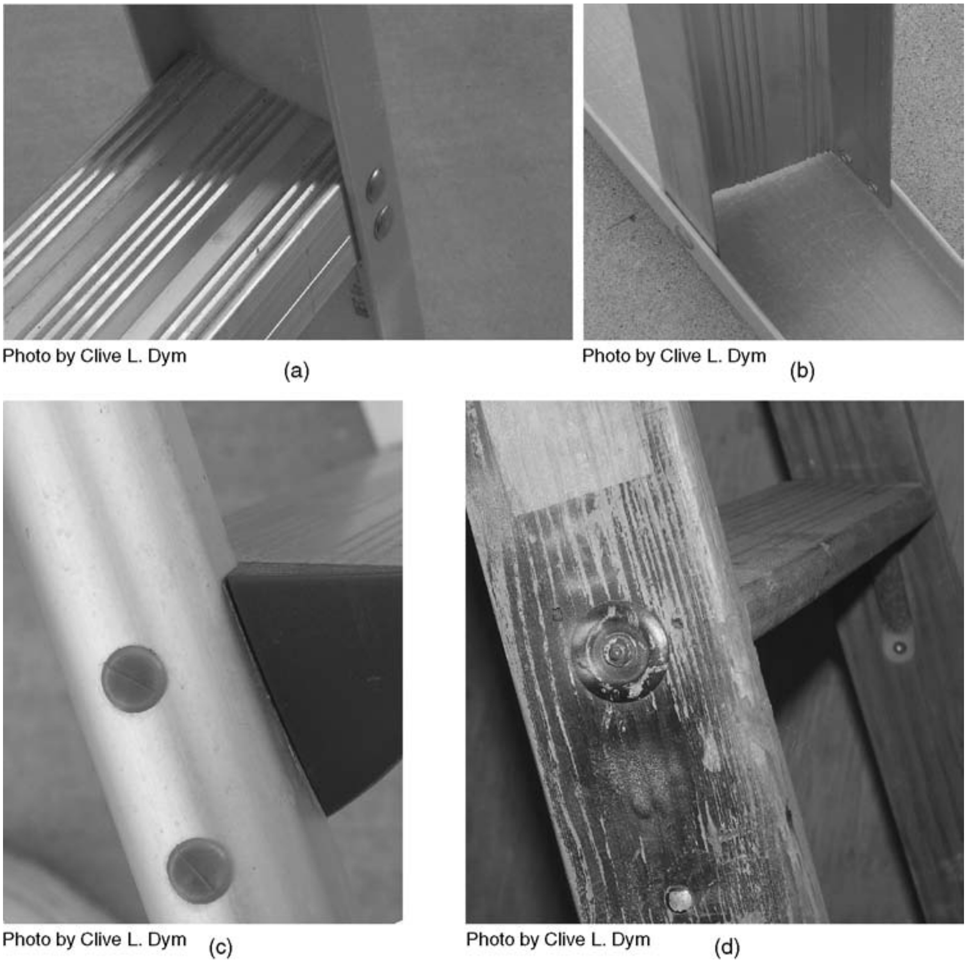



**Figure 12.10** Connecting ladder rungs to ladder frames. (a, b) Top and bottom views of how metal rungs are attached to a ladder's fiberglass frame; note the gap between the top surface of the rung and the frame, so the support is neither simple nor fixed. (c) On this metal ladder the connections go through the frame into the hollow (curved) box that forms the rung; it is also an intermediate support. (d) This very old wooden stepladder has solid rungs that are bolted to the frame, but there are extra (partially visible) supports that bring the connection closer to fixed or clamped.

### 12.4.2 Design criteria

What are our design criteria, that is, against what requirements do we assess the performance of our designs? In part, that depends on both our objectives and our constraints. We identified some top-level objectives during conceptual design:

- *light weight*, that is, minimize the mass of material used in the ladder; and
- *inexpensive*, that is, minimize the cost.

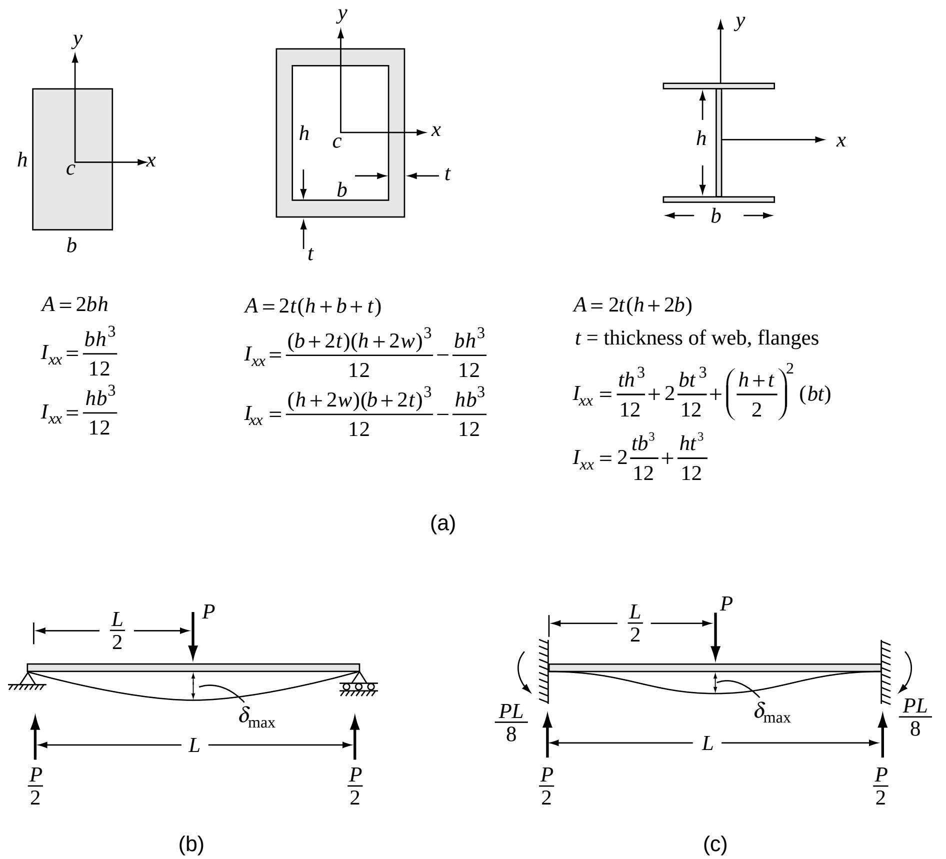



**Figure 12.11** Some aspects of elementary beam models: (a) rectangular, hollow box, I-beam and channel cross-sections, including thicknesses ($h$) and second moments ($I$); (b) a beam with *simple supports* at both ends; and (c) a beam with *fixed supports* at both ends.

We should keep in mind that the relative importance of these two objectives will vary with our client and application: Mass is far more important for a ladder to be used on a space vehicle, while a "big box" retailer is more likely to worry about cost. But there are two other design aspects that must be considered and that can be categorized either as objectives or as constraints. Those two issues derive from not wanting the rung to break or fail when someone stands on it, and not wanting the rung to deflect too much lest that someone feel uncomfortable. We need to specify what it means to require that the rung "not break or fail" and "not deflect too much." And we need to specify whether these are both (possibly competing) objectives, constraints, or some combination.

A material breaks or fails when any one of three *failure strengths* is exceeded. When we determine values of the design variables such that the rung's bending stress does not exceed specified failure strengths, we are *designing for strength*. The three failure strengths are values of the stresses at which a material fails under, respectively, a *tensile* stress, a *bending* test, or a *tensile test that produces permanent deformation*. These three failure strengths are materials properties that have been measured and tabulated for most materials. As a result, this aspect of our design problem is in part a *materials selection* problem. We can generally lump the three modes of failure together and refer to the minimum of the three for a given material as the failure strength of interest, $\sigma_f$. Since material properties are largely established through laboratory testing and experience, our degree of confidence varies with the material. We reflect that variation of confidence by stating that the failure strength should be divided by a safety factor $S$, with $S$ being as low as 1.2 for well-understood materials and as high as 5 for materials for which the properties are not as well established. (Of course, other uncertainties can be incorporated into $S$.) Then the strength requirement would be expressed in terms of the bending stress as $\sigma \leq \sigma_f/S$. Failure strengths typically have values on the order of a megapascal, $1\,\text{MPa} = 10^6\,\text{Pa}$, where the pascal (Pa) is the SI unit of stress defined as $1\ \text{Pa} = 1\,\text{N/m}^2$.

A rung deflects too much when a specified maximum deflection is exceeded. When we determine values of the design variables such that the rung's midpoint deflection does not exceed specified deflection limits, we are *designing for stiffness*. The specified upper limit generally follows from ergonomic considerations: We don't want a ladder to feel wobbly when we are standing on a rung. Thus, codes or standards often specify a maximum deflection $\delta_{max}$ as a fraction of the rung's length, $L$. The deflection requirement would then be expressed in terms of the deflection as $\delta \leq \delta_{max} \leq C_d L$, where $C_f$ is a very small number, say $C_d \sim L/100$.

We now address one last loose (design) end: Do we choose pairs of constants, $C_\delta$ and $C_\sigma$, to correspond to beams with simple supports or to beams with fixed supports? Experience suggests that the ends of a step would be most accurately modeled as fixed or clamped. However, since the side rails of the frame are not truly rigid, there will always be a (very) small amount of rotation at the rung's ends (see Figure 12.10). Thus, we will model the rung as a beam on simple supports, knowing it will be a more flexible (and thus more conservative) model that will overpredict both the stress in and the deflection of the step [***Assume***]. As a result, our final design will bend less and carry bigger loads than our model predicts. The values of the constants for simple supports, $C_\delta = 48$ and $C_\sigma = 4$, are found from exact solutions for the deflection and the bending stress of a simple beam carrying a vertical load $P$ at its midpoint.

To summarize, we have four objectives to achieve: We want to minimize both the mass and the cost, subject to both the strength constraint and the stiffness (or deflection) constraint. There are several ways to proceed. We could simply look at how the mass and the cost vary with different materials and then assess designs for both strength and stiffness. If we were experienced structural designers or our intuition were sufficiently well developed, we might note that the stiffness constraint is usually much more severe than the strength constraint. That is, if the stiffness constraint ($\delta \leq \delta_{max} = C_f L$) is met, it is rather unlikely that the strength constraint ($\sigma \leq \sigma_f/S$) will be violated. If such a case emerged, we might have to revise our objectives for cost and our constraint on stiffness, and then check again whether the strength is sufficient. We also want to ensure that we get a reasonable result for the step thickness. For example, polymer foams may be superior in both cost and stiffness, but the final step thickness may be 0.5 m, which is clearly impractical for a ladder. (Such a large thickness would also violate the assumptions underlying the beam model whose results are given in eqs. (12.63) and (12.64). [***Validate***])

## 12.5 PRELIMINARY DESIGN OF A LADDER RUNG

We now undertake some elements of preliminary design. In the "real" preliminary design of a ladder we would consider beams of various cross-sections and likely make some estimates of which shapes are likely to be more efficient when made of different materials. We might then choose one shape for further development, along with a range or set of materials. Then, in "real" detailed design we would refine that design by working to optimize it, making it as light and cheap as possible. We would also decide how to attach the rungs to the ladder frame (e.g., with rivets, welds, or bolts) and then "size" and fix the locations of those attachments. In our case, we will use preliminary design to illustrate and contrast materials selection when designing for strength and when designing for deflection (or stiffness). In our detailed design we will optimize rung designs to achieve minimum mass and minimum cost.

### 12.5.1 Preliminary Design Considerations for a Ladder Rung

With both failure and deflection criteria defined, and both loose ends tied up, our design problem is that we want the rung's midpoint deflection $\delta$ and its maximum bending stress $\sigma$ to satisfy

$$\frac{\delta}{L} = \left(\frac{1}{L}\right)\frac{PL^3}{48EI} \leq \frac{\delta_{max}}{L} = C_f \tag{12.65}$$

and

$$\sigma = \frac{PL}{8hI} \leq \frac{\sigma_f}{S} \tag{12.66}$$

and where $P$ represents the combined weight of someone standing on the ladder and of the package that person is carrying, that is, $P = W_p + W_w$. An important question is, do we treat these as inequalities (objectives) or do we adopt the equal signs (constraints) for both the deflection and the stress? The answer is that we cannot treat both as constraints. While there are nominally three design variables ($E$, $I$, $h$), $I$ and $h$ are so strongly related that they are effectively a single variable. Still more important, $E$ and $I$ (and $h$) are not truly independent variables. In fact, as we pointed out in the discussion that follows eq. (12.64), the material property ($E$) and the geometric properties ($I$ or $h$ and $L$) are incorporated into a single effective stiffness, namely, $k_{eff} = 48EI/L^3$. The implication of this is that we can design for strength or we can design for stiffness, but we cannot design for both at once. Thus, we can choose to minimize the deflection, in which case we are designing for stiffness and we must then check to ensure that the corresponding bending stress is below the failure criterion. To design for stiffness, we start by equating the bending stress to the failure strength, after which we calculate the corresponding deflection and assess whether we can (or not) accept that value.

As noted, the rung's effective stiffness depends on both material and geometric properties. In many structural design problems, the material is chosen or specified in advance. In that case, the design variable that remains is the cross-sectional area of the rung as represented by its second moment, $I$. As shown in Figure 12.11(a), the second moment $I$ and the rung's thickness $h$ can be used to model a wide variety of shapes, including rectangular cross-sections, I-beams, and channel sections. The reasons that I-beams, channels, and similar such sections are widely used are that they are more efficient than rectangular cross-sections in that they sustain higher stresses per unit weight, and that modern material processing capabilities make it easy to manufacture such shapes in great volume. In fact, if we look again at Figure 12.10, we see that only the wooden ladder rung has a complete rectangular cross-section. However, knowing that our results will perhaps be unrealistic, we will limit our exploration of this aspect of the design space by assuming that the step has a rectangular cross-section, with width $b$ [***Assume***]. In this case, then, $I = bh^3/12$, and our current set of design variables has changed from $E$, $I$, and $h$ to $E$, $b$, and $h$.

Finally, we will also assume that the rung width $b$ is constrained, as in fact it is. For example, the *American National Standard for Ladders—Wood Safety Requirements* (published by The American National Standards Institute) stipulates fixed rung widths for a variety of wooden stepladders. Thus, even though we are not constraining our design to wooden rungs, we will assume that the width $b$ is a specified quantity [***Assume***]. As a result, we now actually have only two design variables, $E$ and $h$.

### 12.5.2 Preliminary Design of a Ladder Rung for Stiffness

We now formulate the first of two different preliminary design problems for rungs of rectangular cross-sections and fixed widths, namely, we *design for stiffness* by constraining the deflection. We do that because we don't know in advance what stiffness is required to achieve a specified deflection, although we do know (and specify) the limiting value that we place on that deflection. We might also think of design for stiffness as *design for deflection*. In eq. (12.65) we see that $C_\delta = PL^2/48EI$, which then yields

$$\frac{PL^2}{4Ebh^3} = C_f \tag{12.67}$$

Equation (12.67) can be solved for the thickness design variable, $h$:

$$h = \left(\frac{PL^2}{4EbC_f}\right)^{1/3} \tag{12.68}$$

Equation (12.68) determines the rung thickness and its value clearly depends on the given values of $P$, $L$, $b$, and $C_f$, as well as on the value of the modulus $E$, which is as yet unspecified. Typically we would have a range of materials in mind (e.g., aluminum, steel, wood, or a composite for a ladder), and we would calculate the corresponding thicknesses of our rung accordingly. However, we have to ensure that the step does not fail, and thus by substituting eq. (12.68) into eq. (12.67) we would check that

$$\sigma = \frac{PL}{8hI} = \frac{3PL}{2bh^2} = \left(\frac{54PC_f^2}{bL}\right)^{1/3} E^{2/3} \leq \frac{\sigma_f}{S} \tag{12.69}$$

**TABLE 12.5 Possible designs of the rung of a ladder for stiffness. The design load (i.e., the weight supported) is $P = 1350$ N; the rung length $L = 350$ mm; the rung width $b = 75$ mm; the safety factor is $S = 1.5$; and the constraint constant $C_d = 0.01$. Note that two of the materials, aluminum and wood, violate our strength constraint that $\sigma/(\sigma_f/S) \leq 1$**

| Material | $E$ (GPa) | $h$ (mm) | $\sigma/(\sigma_f/S)$ |
|---|---|---|---|
| Aluminum | 70 | 9.2 | 1.51 |
| Steel | 212 | 6.4 | 0.63 |
| Wood | 9 | 18.3 | 1.06 |
| CFRP | 110 | 7.9 | 0.90 |

Or, with the stress cast in a dimensionless ratio,

$$\frac{\sigma}{\sigma_f/S} = \left(\frac{54PC_f^2}{bL}\right)^{1/3} \frac{E^{2/3}}{\sigma_f/S} \leq 1 \tag{12.70}$$

In Table 12.5 we show some thickness and modulus values of possible rungs designed to meet the stiffness constraint. We note that the thicknesses all seem reasonably small, but that the ratio of the stress to the failure stress, $\sigma/(\sigma_f/S)$, is not always less than 1. Our intuition that the stiffness constraint would be more severe than the strength constraint turned out to be false for two of the materials, although the wood rung is a very near miss, even with a small safety factor. [***Verify***] A different kind of wood might have produced a more satisfying result. However, these results confirm one of the basic reasons for making a model, that is, to (numerically) check on our intuition.

### 12.5.3 Preliminary Design of a Ladder Rung for Strength

We *design for strength* by constraining the stress to ensure that failure does not occur. We then have to calculate the corresponding deflection and decide whether we can accept that deflection. Thus, the statement of designing for strength begins with eq. (12.68), $\sigma = PL/8hI = \sigma_f/S$, which means constraining the stress such that

$$\sigma = \frac{PL}{8hI} = \frac{3PL}{2bh^2} = \frac{\sigma_f}{S} \tag{12.71}$$

We now solve for the thickness design variable, $h$, using the constraint (12.71):

$$h = \left(\frac{3SPL}{2b\sigma_f}\right)^{1/2} \tag{12.72}$$

Then the deflection that corresponds to this strength design is formulated by substituting eq. (12.72) into eq. (12.65):

$$\frac{\delta}{L} = \left(\frac{bL}{54P}\right)^{1/2} \frac{\left(\sigma_f/S\right)^{3/2}}{E} \tag{12.73}$$

Note, too, that here we have cast the deflection in a dimensionless ratio.

**TABLE 12.6 Possible designs of the rung of a ladder for strength. The design load (i.e., the weight supported) is $P = 1350$ N; the safety factor is $S = 1.5$; the rung length is $L = 350$ mm; and the rung width $b = 75$ m. Note that two of the materials, steel and CFRP, violate our stiffness constraint that $\delta/L \leq 0.01$**

| Material | $E$ (GPa) | $\sigma_f$ (MPa) | $h$ (mm) | $\delta L$ |
|---|---|---|---|---|
| Aluminum | 70 | 110 | 11.4 | 0.005 |
| Steel | 212 | 550 | 5.1 | 0.020 |
| Wood | 9 | 40 | 18.8 | 0.009 |
| CFRP | 110 | 250 | 7.5 | 0.012 |

In Table 12.6 we show some results for strength design for the same four materials listed in Table 12.5. Here, too, the thicknesses all seem reasonably small, but the ratio of the midpoint deflection to the rung length, $\delta/L$, is sometimes—but not always—larger than the 0.01 limit prescribed in the stiffness design. The two materials that failed the strength constraint in the stiffness design, aluminum and wood, here pass the stiffness constraint in the strength design, and vice versa for the other two, steel and CFRP. We were right only half of the time with our assumption about stiffness being a more severe constraint than strength. [***Verify***]

## 12.6 CLOSING REMARKS ON MATHEMATICS, PHYSICS, AND DESIGN

We have seen that preliminary design requires both careful mathematical modeling and appropriately focused research to obtain relevant data. We have examined two cases of mathematical modeling and their results, but designing beams and doing research to identify appropriate materials are both eminently practical skills in their own right. Further, the lessons regarding dimensions, scaling, simplifying assumptions, and how a model answers only the questions asked of it are lessons that can be applied to almost all modeling (and design) efforts.

One of the reasons that engineering programs emphasize engineering science content is a reflection of the extent to which such modeling and related research activities must be performed to do good engineering.

## 12.7 NOTES

*Section 12.1*: The discussion of mathematical modeling has its roots in Dym (2004) and Dym (2007).

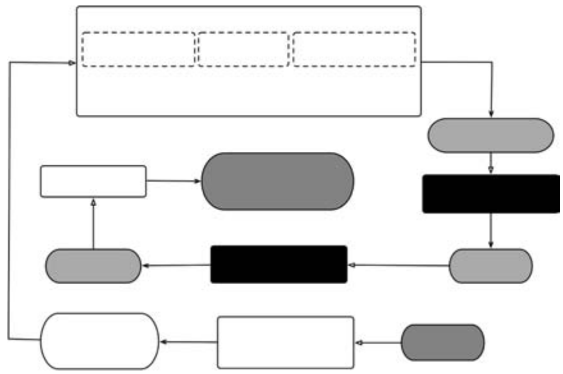
*Section 12.2*: The modeling of beams is taken from Dym (1997), while the approach to material selection is due to Ashby (1999).

*Section 12.3*: Again, the approach to material selection, especially the use of materials selection charts, is based on Ashby (1999).

CHAPTER 13

# ENGINEERING ECONOMICS IN DESIGN

*How much is this going to cost?*



DESIGNS ALMOST always have to meet cost-related or other economic targets. Therefore, it is essential that we understand how to estimate and manage the costs associated with our design work. In this chapter, we introduce two basic concepts of engineering economics in the context of the design process, after which we close with some remarks about the field of engineering economics.

## 13.1 COST ESTIMATION: HOW MUCH DOES THIS PARTICULAR DESIGN COST?

How do we estimate the costs of a design, both as we make it and then over its planned life? In practice, cost estimation is a complex business that requires skill and experience. However, there are several ways that we can break out the cost structure of a device that we are designing. The simplest, conceptually, is to estimate labor, materials, and overhead costs. This simple statement ignores profits, and it masks the complexity of the full cost-related details of all but the simplest of artifacts. Nevertheless, we will limit our discussion to describing only the principal elements that make up the cost categories listed above, since they allow designers to begin to grasp the economics of a design.

Before turning to particular cost categories, it is useful to remember that "low cost" is very likely to be among the objectives of a designed artifact. Even when "money is no object," we will prefer a less expensive solution to a more costly one if both are equal in all other respects. This means that a designer should understand fundamental elements of costing out a design, even if only to answer the inevitable question, "How much?"

### 13.1.1 Labor, Materials, and Overhead Costs

Costs are often broken up into the categories of labor, material, and overhead costs. *Labor costs* include payments to the employees who build the designed device, as well as to support personnel who perform necessary but often invisible tasks such as taking and filling orders, packaging, and shipping the device. Labor costs also include a variety of *indirect costs* that are less evident because they are generally not paid directly to employees. These indirect costs are sometimes called *fringe benefits* and include health and life insurance, retirement benefits, employers' contributions to Social Security, and other mandated payroll taxes. We should not overlook these indirect costs of labor when we are estimating the cost of a design because in many companies they are as much as 50% of direct labor payments or wages. Of course, we need to carefully study the activities required to produce an instance of our design in order to estimate the labor costs. It is easy to overstate how much time is involved, both because we may be building the very first instance, and because we may also be learning how to use new tools and technology. We should review our labor estimates with professionals experienced in production before we consider them accurate. Having said this, a simple starting point for estimating costs is to keep good records or the activities needed to build our design's prototype. Similarly, if we have developed a process router to describe the assembly tasks and their sequence, we can use that router to estimate minimum and typical times to construct a mechanical system.

*Materials* include those items and inputs directly used in building the device, along with intermediate materials and inventories that are consumed in the manufacturing process. A key tool for estimating the materials cost of an artifact is the *bill of materials* (BOM), the list of all of the parts in our design, including the quantities of each part required for complete assembly. The BOM is particularly useful since it is usually developed directly from the assembly drawings, and so it reflects our final design intentions. We might think of the BOM as being that part of a recipe that specifies all of the ingredients that we need, as well as the exact quantities needed to make a specified lot size (i.e., the number of the designed devices we want to make). Figure 13.1 shows the BOM for a screwdriver. Notice that the list of parts is included directly in the drawing in this case. For a more complex part, we might have a separate BOM, both for the clarity of the drawing and to make it easier for purchasing agents to collect similar parts used to make a variety of devices.

The BOM lists all the materials used to make the device. As such, it matters greatly whether our design starts with raw materials or with materials that have already been worked by our suppliers. In the case of the screwdriver, we might decide to "outsource" the handle, in which case our BOM would include a prepared handle rather than a stock piece of cast acrylic that must be machined. Such outsourcing affects the cost of both the materials and the labor needed to transform raw materials into finished products.

Materials costs can often be reduced significantly by using commercial off-the-shelf materials rather than making our own. This is because outside vendors have the machinery and expertise to make very large numbers of parts for a lot of customers. It is a good habit of thought to look at parts catalogs and other listings of premade components and parts.

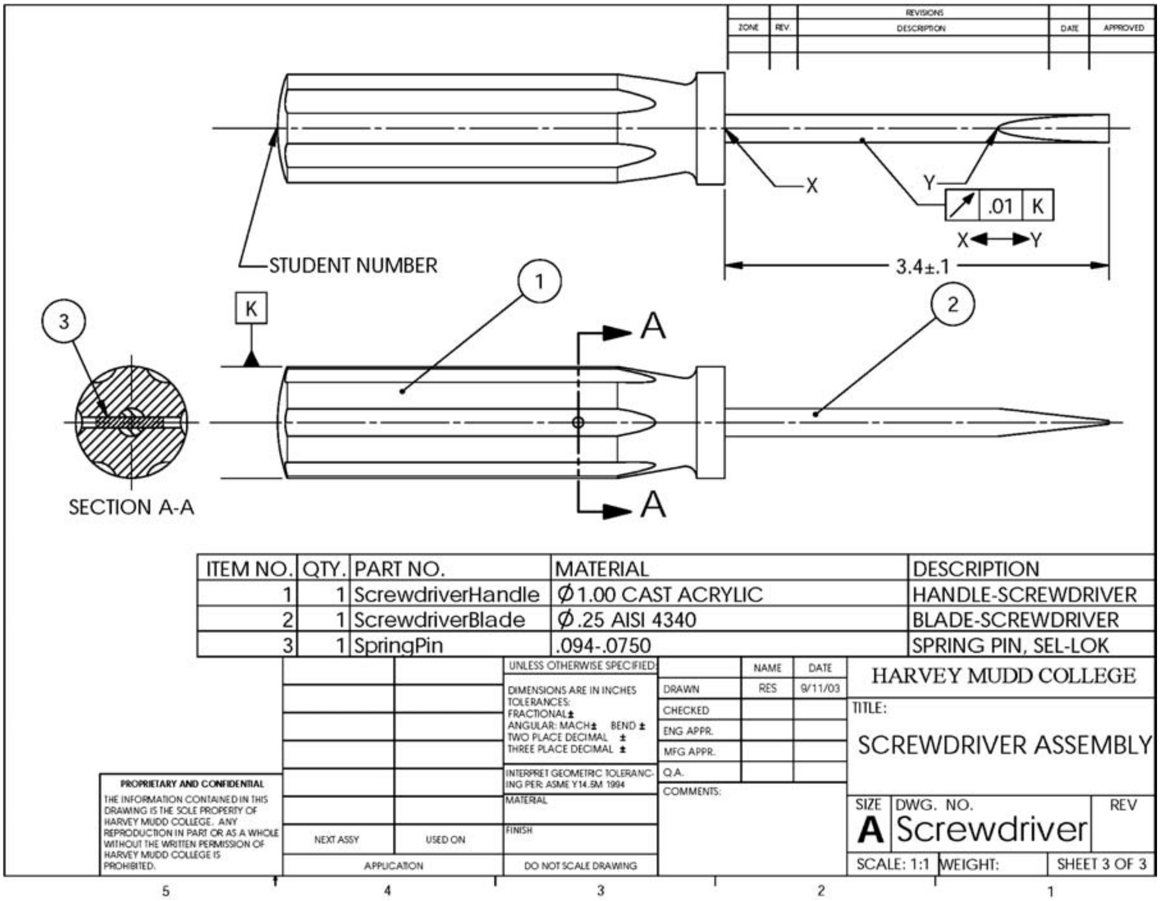



**Figure 13.1** The bill of materials (BOM) for a screwdriver, listing all of the materials needed to fabricate the part. In this case, the BOM is built into a drawing, but it could be given as a separate list. (Drawing courtesy of R. Erik Spjut.)

The costs incurred by a manufacturer that cannot be directly assigned to a single product are termed *overhead*. If, for example, a company makes a product in a factory that also produces 20 other products, the cost of the building, the machines, the janitorial staff, the electricity, and so on, must somehow be shared or distributed among all of the 21 items. If the company ignored these overhead costs in setting its product's prices, it would soon find itself unable to pay for the building and the services necessary to maintain it. Other elements of overhead include the salaries of executives, who are presumably using some share of their time to supervise each of the company's activities, as well as the costs of needed business functions such as accounting, billing, and advertising. While there are accounting standards that define cost categories and their attributes, precise estimates of overhead costs vary greatly with the structure and practices of the company in question. One company may have only a small number of products and a very lean organization, with most of its costs directly attributed to the products made and sold, and only a small percentage allocated to overhead. In other organizations the overhead can be equal to or greater than the labor costs that are directly assignable to one or more products. Estimating the costs of producing a design requires careful consultation with clients or their suppliers.

Cost estimates produced during the conceptual stage of a design project are often quite inaccurate when compared with those made for detailed designs. In heavy construction projects, for example, an accuracy of $\pm 35\%$ is considered acceptable for initial estimates. However, this wide range should not be used to justify sloppy or casual cost estimation.

In practice, each engineering discipline has its own approaches to cost estimating that are captured by general guidelines (i.e., *rules of thumb* or *heuristics*) that are useful at the conceptual design stage. In civil engineering, for example, the *R. S. Means Cost Guide* provides cost estimates per square foot for the various elements in different kinds of construction projects. The *Richardson's Manual* offers similar information for chemical plant and petroleum refinery projects. Costs per square inch may be more relevant for certain printed circuit board designs. As designers, we should consult with experienced professionals in the appropriate engineering discipline to estimate costs successfully, even when we are making "only" conceptual design choices.

### 13.1.2 Economies of Scale: Do We Make It or Buy It?

Both labor and materials are subject to *economies of scale*, that is, the idea that we can reduce the *unit production cost* of a design by making a lot of copies, rather than just making a few "originals." Consider the difference between the cost of making one soda can in a machine shop and the unit costs to the bottling companies that make millions of cans daily using highly engineered tooling and materials. High volume production allows companies to distribute the costs of specialized and innovative technologies over a lot of units, thus lowering the cost per unit produced. This can have a profound effect on our designs.

Economies of scale influence our design choices in several ways. First, they require that the designer know the nature and size of the market in which the design will be used: If there are few sales, then production costs must be shared by a small number of items, which raises the (estimated) unit cost. Even more significant for many products, economies of scale affect our selection of the components that go into our designs. In many cases, clients will insist that we use standard components that are produced and sold in high volumes whenever possible, to ensure both lower costs and adequate supplies of replacement parts.

Because of economies of scale and the value of specialized knowledge, it is almost always the case that a design engineer will choose a fully functional component produced in large volumes over a unique made-in-house component.

### 13.1.3 The Cost of Design and the Cost of the Designed Device

There is an important distinction between the cost of designing, prototyping, and testing a product and the cost of the product after manufacturing and distribution. Consider a new microprocessor. A company may spend many millions of dollars to design and develop a leading-edge chip, with the intention that it will initially sell the final products for tens of dollars, and eventually for a dollar or less. In other cases, the cost of design is a relatively small part of the final project cost, such as a building a dam or skyscraper. Notwithstanding such differences, most clients expect that a design team will correctly estimate its own costs and budget them accurately. Thus, even when costing out the design activity, an effective design team has to understand and control its costs accurately.

It should also be noted that while costing is an important element in the *profitability* of a design, it is generally *not* a key factor in the *pricing* of the artifact. This seeming contradiction can be easily explained by noting that gross profits (i.e., profits before taxes and other considerations) are simply the net of revenues minus costs. As such, costs are an important element in the profit equation. Revenues, on the other hand, are determined by the price charged for an item multiplied by the number of items sold. For most profit-maximizing firms, prices are not set on the basis of costs, but rather in terms of what the market is willing to pay. Consider our leading edge microprocessor. It may command a price of hundreds of dollars when initially introduced. However, the costs of making each chip, amortized over the full production run, are well below that. The materials in the chip may cost pennies, the labor per chip is essentially negligible, and even the costs of research and innovative production lines are relatively modest when shared among the many millions of chips ultimately sold. Distribution costs are clearly no different for new chips than for older ones. But since the company anticipates great demand for its new leading edge processors, it will set their price high: There are customers willing to pay a premium to get them. Indeed, the responsibilities of marketing professionals on a design team usually include identifying design attributes that make consumers willing to pay a high price for a new product design.

## 13.2 THE TIME VALUE OF MONEY

Beyond the immediate costs of making a new designed product, we also have to consider the economic consequences of our design choices over the full lifetime of the designed artifact. This is made more difficult because the value of money changes depending on when we receive or pay it out. If someone were to offer us $100 today or $100 a year from now, we would prefer to take the money *now*. Having the money sooner offers us a number of advantages. We can invest or otherwise use the money during the intervening year; we also eliminate the risk that the money might not be available next year; and we don't have to worry that inflation will shrink the purchasing power of that $100 during the next year. This simple example highlights an important concept, the *time value of money*: Money obtained sooner is more valuable than money obtained later, and money spent sooner is more costly than money spent later.

The time value of money captures the effects of both *opportunity costs* and *risk*. An *opportunity cost* is a measure of how much the deferred money could have earned in the intervening time. The *risk* captures the chance that the money will be worth less (because of inflation), and that changed circumstances could make the money unavailable during the intervening time. Economists and financial professionals bundle together the extent of these risks and lost opportunities in a single *discount rate* that acts much like an interest rate on a savings account or credit card: The interest rate on a savings account measures how much a bank is willing to pay for the privilege of using our money in the coming year. An interest rate on a credit card measures how much we must pay the card issuer in the future for the privilege of using their money now. Just as these rates vary based on the credit ratings of customers, discount rates will vary depending on what other uses and risks are faced by investors. Interest calculations on credit cards and bank accounts typically show dollar amounts increasing from a given time onward or forward. On the other hand, discount rates typically work in reverse, showing the value today of money that will be available or required at some point in the future.

How do we distinguish between "\$100 today" and "\$100 a year from now" in a rational and consistent way? For a given a set of future financial events or cash flows (both toward us and away from us), we use a given discount rate to translate all of our financial events into a common time frame, either the current time or a future one. Consider again our choice of getting \$100 today or \$100 next year. If the annual discount rate is 10%, then we would expect to need \$110 dollars in a year in order compensate us for not having \$100 today (or to purchase then what we could buy today with \$100). We would be getting the same value if we accept \$100 now or \$110 a year from now. We can write this up more generally as

$$V_F = (1 + r)V_P \tag{13.1}$$

Here $V_F$ is the future value of the money (a year from now), $V_P$ is the present value (at this time), and $r$ is the discount rate expressed in percent.

We can also work this the other way, asking how much we need today to have the equivalent of \$100 a year from now. In this case, $V_F = \$100$, $r$ is 0.1, and we now invert eq. (13.1):

$$V_P = \frac{V_F}{(1 + r)} = \frac{\$100}{1.10} = \$90.91 \tag{13.2}$$

Taking the \$100 next year is equivalent to getting and investing about \$91 today, given our 10% discount rate. We can carry out these discounting calculations as far into the future as we want. The principle is the same. Thus, \$100 promised for two years from now would be worth even less than the \$100 promised for next year, since it doesn't account for our being able to use the \$100 for two years, or our being able to use the \$10 earned in the first year.

Economists have developed standard approaches to discounting money and to determining the present value of future dollars, whether costs or benefits, and vice versa. Application of these formulas can become quite involved when inflation or unusual timing issues are involved, but virtually all such analyses are based on the relationship

$$V_P = \frac{V_F}{(1 + r)^t} \tag{13.3}$$

In eq. (13.3) $V_P$ is the present value of the costs or benefits, $V_F$ is the future value, $r$ is the discount rate, and $t$ is the time period measured in years over which a cost is incurred or a benefit is realized. Consider the case where we are interested in the present value of a payment of \$100 after three years from now. In this case, $V_F = \$100$, $t = 3$, and $r = 0.10$. If substitute these values into eq. (13.2), we find that $V_P = \$100/(1 + 0.10)^3 = \$75.13$. Thus, our payment of \$100, deferred to the third year, is equivalent to an immediate payment of some \$75.

We can also use eq. (13.3) to look at a mixture of costs over several time periods as well. Consider the case where we have designed a system with an initial cost of \$1000, annual maintenance costs of \$100, and an overhaul cost of \$300 in the third year. Suppose further that the system is intended to last for five years. We can take the present value of each of these costs and sum them to get the total present value of the cost stream for the system:

$$V_P = V_{P/initial} + V_{P/annual} + V_{P/overhaul} \tag{13.4}$$

Since the initial cost occurs at the outset, we don't need to discount that cost at all: \$1000 today has a present value of \$1000. We can determine the present value of annual maintenance costs for each of the five years using eq. (13.2) and sum them together, while rounding to the nearest dollar:

$$V_{P/annual} = \sum_{t=1}^{5} \frac{\$100}{(1+0.10)^t} \cong \$(91+83+75+68+62) = \$379 \tag{13.5}$$

Finally, we use eq. (13.2) to discount the \$300 overhaul cost in the third year:

$$V_F = (1+r)V_{P/overhaul} = \frac{\$300}{(1+0.10)^3} = \$225 \tag{13.6}$$

Thus, the sum of present values of the initial cost (\$1000), the annual maintenance cost (\$379), and the third-year overhaul (\$225) leads to a present value of the total lifetime cost of \$1604.

This ability to compare costs over the life of a system by putting them into a single time frame is quite important for choosing between designs. Consider an alternative design to the example we just gave, in which the initial cost is \$1500, the annual maintenance is only \$25, and no overhaul is required. If we apply the same approach as above, we will find the present value of the lifetime cost of this second design to be \$1595. Thus, in this case, a design that costs 50% more at the outset turns out to have lower lifetime costs than a seemingly much cheaper design because it has small maintenance and overhaul costs. In practice, the difference between \$1604 and \$1595 is so small that we would probably consider the both designs as economically equivalent. This means that if our client asked for low cost, we would ask whether they want low initial cost, low maintenance costs, or low life cycle costs. Discounting helps us answer such questions and select better designs, because we can properly translate all of the anticipated costs to a common time frame.

What does all this mean for us as designers? Firstly, we need to realize that design decisions and choices made today will translate into financial events that will occur at different times in the future. Our design choices will lead users to incur costs in the future, such as maintenance expenses or higher energy consumption; some of these costs may dominate the initial production cost. Purchasers and users of our design will incur these costs over a time horizon, including the initial purchase price, operating costs, maintenance costs, and eventually disposal and replacement costs. So when our client tells us that they consider "low cost" to be an important objective, we need to question them to find out what "low cost" means to them.

## 13.3 CLOSING CONSIDERATIONS ON ENGINEERING AND ECONOMICS

Engineering and economics have been closely linked for almost as long as the two fields have existed. Indeed, economists recognize that engineers were the developers of a number of important elements of economic theory. For example, *utility theory* and *price discrimination* were both first articulated by the 19th century engineer Jules Dupuit, and *location theory* was developed by a civil engineer named Arthur Wellington. Wellington is also credited with the definition of engineering as "the art of doing that well with one which any bungler can do with two." This link of engineering to economics should come as no surprise to the designer, since it is a rare project for which money really is "no object." Virtually every engineering decision has economic consequences.

The (formal) field of *engineering economics* is concerned with understanding the economic or financial implications of engineering decisions, including choosing among alternatives (e.g., *cost–benefit analysis*), deciding if or when to replace machines or other systems (*replacement analysis*), and predicting the full costs of devices over the period of time that they will be owned and used (*life cycle analysis*). These topics can easily fill entire courses in an engineering curriculum and are well beyond our current scope. We have covered two engineering economics topics—the *time value of money* and *cost estimation*—that are very important to design and should help design and engineering teams make good decisions.

While economic factors are extremely important in design, we should keep in mind that engineering economics is another form of engineering modeling, just like prototyping or simulation; it should not be the only factor influencing a designer. There are other "quasi-economic" modeling approaches that can be used to evaluate the effect of designs, including the so-called "triple bottom line," which is used to estimate the economic, social, and environmental consequences of a design. In Chapter 14, we will look at how the life cycle of a product can be analyzed in terms of environmental effects, which some people think of as a type of noneconomic cost.

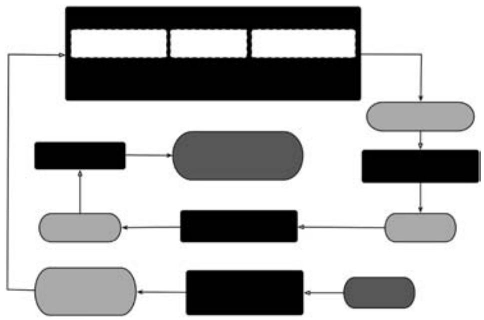
## 13.4 NOTES

*Section 13.2:* There are many excellent texts in engineering economics that can be used to go further with these topics. The cost estimating material draws upon class notes prepared by our colleague, Donald S. Remer for his course on cost estimating and on Oberlander (2000). A view of engineering costing that is based more on accounting may be found in Riggs (1994). The relationship between pricing and costs is discussed in Nagle (1987) and Phlips (1985).

CHAPTER 14

# DESIGN FOR PRODUCTION, USE, AND SUSTAINABILITY

*What other factors influence the design process?*



A CENTRAL theme of this book is that engineering design is usually done by teams. Design teams usually include not only engineers, but also manufacturing experts (who may be industrial engineers), marketing and sales professionals, reliability experts, cost accountants, lawyers, and so on. Such teams are concerned with understanding and optimizing the product under development for its *entire life*, including its design, development, manufacturing, marketing, distribution, use, and, eventually, disposal. Concern with all of these areas, coupled with the need to reduce the time to bring a product to market, has led to what is known as *concurrent engineering*. Concurrent engineering means that a multidisciplinary design team works simultaneously and in parallel to design a product, a manufacturing approach, a distribution scheme, user support, maintenance, and ultimately disposal. While it is beyond the scope of this book to explore concurrent engineering in depth, it is important that engineers understand and appreciate these related fields that influence both the initial design and the full life cycle of a product.

Engineers have always sought to realize various desirable attributes to some degree in their designs, embodied in the design process as objectives. This is often referred to as "design for *X*," where *X* is an attribute such as manufacturing, maintainability, reliability, or affordability. Designers and engineers also refer to them with a different name, the *-ilities*, because many of these desirable attributes are expressed as nouns that have an "-ility" suffix. As designers, we can use the idea of the product life cycle to guide us through some of these *X*'s. Since most products are designed to be built, sold, used, and then disposed of, we can look first at design for manufacturing and assembly, and then at design for use, including reliability and maintainability, and lastly, at design for sustainability. We have already considered design for affordability in Chapter 13, where we look at economic aspects of design. All these related concepts can also be considered aspects of *quality*, which we considered when we looked at quality function deployment and the house of quality.

## 14.1 DESIGN FOR PRODUCTION: CAN THIS DESIGN BE MADE?

In many cases, a designed artifact will be produced or manufactured in large quantities. In recent years, companies have come to learn that the design of a product can have an enormous impact on the methods and costs of producing it. Toward this end, globally competitive industries such as the automotive and consumer electronics industries routinely consider how a product is manufactured during the earliest stages of design. A significant driver of this concern is the number of products being manufactured, which may allow for the economies of scale that we discussed in Chapter 13. Further, the time it takes to get a product to the consumer, known as the *time to market*, defines a company's ability to shape a market. Design processes that anticipate manufacturing issues can be key elements in speeding products through to commercial production.

### 14.1.1 Design for Manufacturing (DFM)

*Design for manufacturing* (DFM) is design based on minimizing the costs of production and/or the time to market for a product, while maintaining an appropriate level of quality. The importance of maintaining an appropriate level of quality cannot be overstated because without an assurance of quality, DFM is reduced to simply producing the lowest cost product.

DFM begins with the formation of the design team. In commercial settings, design teams committed to DFM tend to be multidisciplinary, and they include engineers, manufacturing managers, logistics specialists, cost accountants, and marketing and sales professionals. Each brings particular interests and experience to a design project, but all must move beyond their primary expertise to focus on the project itself. In many world-class companies, such multidisciplinary teams have become the de facto standard of the modern design organization.

Manufacturing and design tend to interact iteratively during product development. That is, the design team itself discovers a possible problem in producing a proposed design or learns of an opportunity to reduce production costs or timing, and the team then reconsiders its design. Similarly, a design team may be able to suggest alternative production approaches that lead manufacturing specialists to restructure processes. In order to achieve fruitful and synergistic interaction between the manufacturing and design processes, it is important that DFM be considered in each and every one of the design phases, including the early conceptual design stages.

A basic methodology for DFM consists of six steps:

1. estimate the manufacturing costs for a given design alternative;
2. reduce the costs of components;
3. reduce the costs of assembly;
4. reduce the costs of supporting production;
5. consider the effects of DFM on other objectives; and
6. if the results are not acceptable, revise the design once again.

This approach clearly depends upon an understanding of all the objectives of the design; otherwise the iteration called for in Step 6 cannot occur meaningfully. An understanding of the economics of production is also required. This topic is usually taught in industrial engineering courses, or courses in manufacturing. In addition to these areas, however, there are engineering and process decisions made by the design team that can directly influence the cost of producing a product. Some processes for shaping and forming metal, for example, cost much more than others and are called for only to meet particular engineering needs. Similarly, some types of electronic circuits can be made with high-volume, high-speed production machines, while others require hand assembly. Some design choices that require higher costs for small production runs may actually be less expensive if the design can also be used for another, higher volume purpose. In each of these instances, we can complete a successful design only by combining deep knowledge of manufacturing techniques with deep design experience.

There are some specific things that design teams can keep in mind when doing DFM. First of all, consulting with experts on manufacturing can often reveal manufacturing techniques that will (or will not) work with your design. Whether they are faculty members at a university, experts at the client's firm, or even retired manufacturing engineers, tapping into their knowledge is tremendously helpful for the designer. Second, production costs can usually be reduced by using commercially available inputs rather than custom parts. The use of off-the-shelf components will also make cost estimation simpler, since there are catalogs listing specifications and prices. Finally, DFM must always be done with the client's objectives in mind—in some fields, ease of manufacturing or even reduced costs may not be uppermost in the client's mind, especially if lives are at stake.

### 14.1.2 Design for Assembly (DFA)

*Design for assembly* is a related, but formally different type of design for *X*. Assembly refers to the way in which the various parts, components, and subsystems are joined, attached, or otherwise grouped together to form the final product. Assembly can be characterized as consisting of a set of processes by which the assembler (1) handles parts or components (i.e., retrieves and positions them appropriately relative to each other), and (2) inserts (or mates or combines) the parts into a finished subsystem or system. For example, assembling a ball-point pen might require that the ink cartridge be inserted into the tube that forms the handgrip, and that caps be attached to each end. This assembly process can be done in a number of ways, and the designer needs to consider approaches that will make it possible for the manufacturer to reduce the costs of assembly while maintaining high quality in the finished product. Clearly then, assembly is a key aspect of manufacturing and must be considered either as part of design for manufacturing or as a separate, yet strongly related design task.

Because of its central place in manufacturing, a great deal of thought has been put into development of guidelines and techniques for making assembly more effective and efficient. Some of the approaches typically considered are:

1. *Limiting the number of components to the fewest that are essential to the working of the finished product*. Among other things, this implies that the designer will differentiate between parts that could be eliminated by combining other parts and those that must be distinct as a matter of necessity. The usual issues for this are to identify:
   - parts that must move relative to one another;
   - parts that must be made of different materials (for strength, for example, or insulation); and
   - parts that must be separated in order for assembly to proceed.
2. *Using standard fasteners and/or integrating fasteners into the product itself*. Using standard fasteners also allows an assembler to develop standard routines for component assembly, including automation. Reducing the number and type of fasteners allows the assembler to construct a product without having to retrieve as many components and parts. The designer should also consider that fasteners tend to induce stress concentrations and may thus cause reliability concerns.
3. *Designing the product to have a base component on which other components can be located* (including designing for the assembly to proceed with as little motion of the base component as possible). This guideline enables an assembler, whether human or machine, to work to a fixed reference point in the assembly process and to minimize the degree to which the assembler must reset reference points.
4. *Designing the product to have components that facilitate retrieval and assembly.* This may include elements of detailed design that, for example, reduce the tendency of parts and subassemblies to become tangled with one another, or designing parts that are symmetric, so that once retrieved they can be assembled without turning to a preferred end or orientation.
5. *Designing the product and its component parts to maximize accessibility, during both manufacturing and subsequent repair and maintenance*. While it is important that the components be efficient in their use of space, the designer must balance this need with the ability of an assembler or repairer to gain access to and manipulate parts, both for initial fabrication and later replacement.

While these guidelines and heuristics represent only a small set of the design considerations that make up design for assembly, they provide a starting point for thinking about both DFA and DFM. A central principle in both is that quality, and specifically functionality, cannot be sacrificed for the sake of manufacturing or assembly.

### 14.1.3 The Bill of Materials and Production

Effective design for manufacturing and assembly requires a deep understanding of production processes, among the most important of which are ways to plan and control inventories. A common inventory planning technique is *materials requirements planning* (MRP). It utilizes assembly drawings to develop a *bill of materials* (BOM) and an assembly chart—sometimes wittily called a "gozinto" chart—that shows the order in which the parts on the BOM are put together. As we saw in Chapter 13, the BOM is a list of all of the parts, including the quantities of each part required to assemble a designed object. We noted that the BOM is also used when estimating some of the costs of producing the designed artifact.

When a company has determined the size and timing of its production schedule, production planners can determine the size and timing of inventory orders. (Most companies now use *just in time* delivery of parts as they try to avoid carrying large inventories of parts that are paid for but not generating revenues until after they are assembled and shipped.) The importance of the assembly drawings and the BOM in managing the production process cannot be overstated. To be effective, the design team must not only follow accurate methods of reporting their design, but the entire organization must be committed to the discipline that any design changes, or *engineering change orders*, will be reported accurately and thoroughly to *all* the affected parties. Most organizations have formal procedures for recording and managing changes to the design and production process.

A final point to note is that manufacturing concerns include both logistics and distribution, so that these elements have also become an important part of design for manufacturing and assembly. Many companies forge links between the suppliers of materials needed to make a product, the fabricators who manufacture that product, and the channels needed to efficiently distribute the finished product. This set of related activities, often referred to as the *supply chain*, requires a designer to understand elements of the entire product life cycle. It is beyond our scope to explore the role of supply chain management in design, except to note that in many industries, successful designers understand not only their own production and manufacturing processes, but also those of their suppliers and their customers.

## 14.2 DESIGN FOR USE: HOW LONG WILL THIS DESIGN WORK?

Design for use ties together the designer–client–user triangle in a powerful way. We often hear the words "user friendly" applied to describe a product, but this is only meaningful if made specific. This specificity may include interfaces between the designed system and the user, skills requirements, physical demands on the user (captured in the field of ergonomics), and even how well the system works relative to unintended applications. Limitations of time and space require us to describe only a few of the rich topics underlying design for use; in this section we consider reliability and maintainability.

Most of us have a personal, visceral understanding of reliability and unreliability as a consequence of our own experience with everyday objects. We say that the family car is unreliable, or that a good friend is very reliable—someone we can count on. While such informal assessments are acceptable in our personal lives, we need greater clarity and accuracy when we are functioning as engineering designers. Thus, we now describe how engineers approach reliability, along with its sister concept, maintainability.

### 14.2.1 Reliability

To an engineer, reliability is defined as "the probability that an item will perform its function under stated conditions of use and maintenance for a stated measure of a variate (time, distance, etc.)." This definition has a number of elements that warrant further comment. The first is that we can properly measure the reliability of a component or system *only* under the assumption that it has been or will be used under some specified conditions. The second point is that the appropriate measure of use of the design, called the *variate*, may be something other than time. For example, the variate for a vehicle might be miles, while for a piece of vibrating machinery the variate might be the number of cycles of operation. Third, we must examine reliability in the context of the functions discussed in Chapter 6, which emphasizes the care we should take in developing and defining the functions that a design must perform. Finally, note that reliability is treated as a probability, and hence can be characterized by a distribution. In mathematical terms, this means that we can express our expectations of how reliable, safe, or successful we expect a product or a system to be in terms of a cumulative distribution function or a probability density function.

In practice, our use of a probabilistic definition enables us to consider reliability in the context of the opposite of success, that is, in terms of *failure*. In other words, we can frame our consideration of reliability in terms of the probability that a unit will fail to perform its functions under stated conditions within a specified window of time. This requires us to consider carefully what we mean by failure. British Standard 4778 defines a failure as "the termination of the ability of an item to perform a required function." This definition, while helpful at some level, does not capture some important subtleties that we, as designers, must keep in mind: It doesn't capture the many kinds of failures that can afflict a complex device or system, their degree of severity, their timing, or their effect on the performance of the overall system.

For example, it is useful to distinguish between *when* a system fails and *how* it fails. If the item fails when in use, the failure can be characterized as an *in-service failure*. If the item fails, but the consequences are not detectable until some other activity takes place, we refer to that as an *incidental failure*. A *catastrophic failure* occurs when a failure of some function is such that the entire system in which the item is embedded fails. For example, if our car breaks down while we're on a trip and needs a repair in order for us to complete the trip, we would call that an in-service failure. An incidental failure might be some part that our favorite mechanic suggests we replace during the routine servicing of our car. A catastrophic, accident-causing failure might follow from the failure of a critical part of the car while we are driving at freeway speeds. Each type of failure has its own consequences for the users of the designed artifact, and so must be considered carefully by designers.

We often specify reliability by using measures such as the mean time between failures (MTBF), or miles per in-service failure, or some other metric. However, we should note that framing the definition of reliability in terms of probabilities gives us some insight into the limitations inherent in such measures. Consider the two failure distributions shown in Figure 14.1. These two reliability probability distributions have the same mean (or average), that is, $MTBF_a = MTBF_b$, but they have very different degrees of dispersion

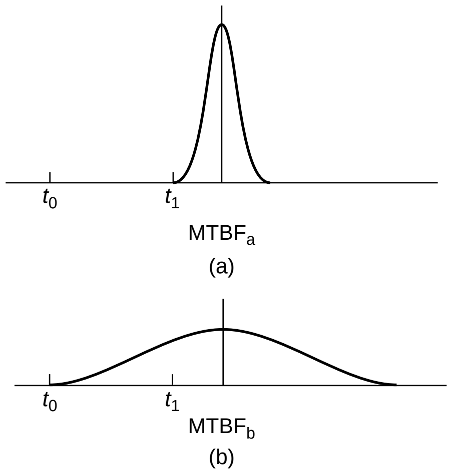



**Figure 14.1** Failure distributions (also called probability density functions) for two different components. Note that both curves have the same value of MTBF, but that the dispersions of possible failures differ markedly. The second design (b) would be viewed as less reliable because more failures would during the early life of the component (i.e., during the time interval $t_0 \leq t \leq t_1$).

(typically measured as the variance or standard deviation) about that mean. If we are not concerned with both mean and variance, we may wind up choosing a design alternative that is seemingly better in terms of MTBF, but much worse in terms of variance. We may even choose a design for which the MTBF is acceptable, but for which the number of early failures is unacceptably high.

One of the most important reliability issues for a designer is how the various parts of the design come together and what the impact is likely to be if any one part does fail. Consider, for example, the conceptual sketch of the *series system* design, shown in Figure 14.2. It is a chain of parts or elements, the failure of any one of which would break the chain, which in turn will cause the system to fail. Just as a chain is no stronger than its weakest link, a series system is no more reliable than its most unreliable part. In fact, the reliability—or probability that the system will function as designed—of a series system whose individual parts have reliability (or probability of successful performance) $R_i(t)$ is given by

$$R_S(t) = R_1(t) \cdot R_2(t) \cdot \cdots \cdot R_n(t)$$

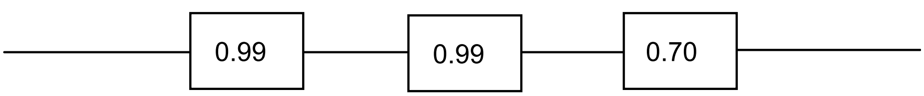



**Figure 14.2** This is a simple example of a *series system*. Each of the elements in the system has a given reliability. The reliability of the system as a whole can be no higher than that of any one of the parts because the failure of any one part will cause the system to cease operating.

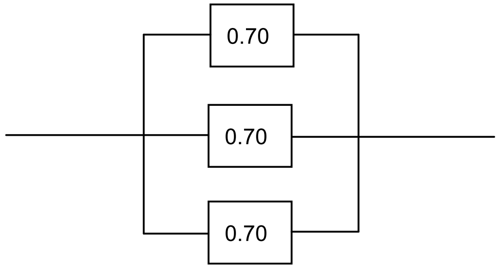



**Figure 14.3** This is a simple example of a *parallel system*. Note that every one of the components must fail in order for the system to cease working. While such a system has high reliability, it is also quite expensive. Most designers seek to incorporate such redundancy when necessary, but look for other solutions wherever possible.

or

$$R_S(t) = \prod_{i=1}^{n} R_i(t) \tag{14.1}$$

Here $R_S(t)$ is the reliability of the entire series system, and $\prod_{i=1}^{n}$ is the product function. We see from eq. (14.1) that the overall reliability of a series system is equal to the product of all of the individual reliabilities of the elements or parts within the system. This means that if any one component has low reliability, such as the proverbial weak link, then the entire system will have low reliability and the chain will likely break.

Designers have long understood that redundancy is important for dealing with the weakest link phenomenon. A *redundant system* is one in which some or all of the parts have backups or replacement parts that can substitute for them in the event of failure. Consider the conceptual sketch of the *parallel system* of three parts or elements shown in Figure 14.3. In this simple case, each of the components must fail in order for the system to fail. The reliability $R_P(t)$ of this entire parallel system is given by

$$R_P(t) = 1 - [(1 - R_1(t)) \cdot (1 - R_2(t)) \cdot \cdots \cdot (1 - R_n(t))]$$

or

$$R_P(t) = 1 - \prod_{i=1}^{n} [1 - R_i(t)] \tag{14.2}$$

We can see from eq. (14.2) that the reliability of this parallel system (i.e., the probability that the parallel system will operate successfully) is now such that if any one of the elements functions, the system will still function.

Parallel systems have obvious advantages in terms of reliability, since all of the redundant or duplicate parts must fail in order for the system to fail. Parallel systems are also more expensive, since many duplicate parts or elements are included only for contingent use, that is, they are used only if another part fails. For this reason, we must carefully weigh the consequences of failure of a part against failure of the system, along with costs attendant to reducing the likelihood of a failure. In most cases designers will opt for some level of redundancy, while allowing other components to stand alone. For example, a car usually has two headlights, in part so that if one fails the car can continue to operate safely at night. The same car will usually have one radio, since its failure is unlikely to be catastrophic. The mathematics of combining series and parallel systems is beyond our scope, but we clearly have to learn and use them to design systems that have any impact on the safety of users. Put simply, redundancy usually increases both reliability and costs.

Designers can consider modes of failure and develop estimates of reliability only if they truly know how components might fail. Such knowledge is gained by performing experiments, analyzing the statistics of prior failures, or by carefully modeling the underlying physical phenomena. Designers lacking deep experience in understanding component failure should consult experienced engineers, other designers, users, and the client in order to ascertain that an appropriate level of reliability is being designed into the system. Often the experience of others allows a designer to answer reliability questions without performing a full set of experiments. For example, the appropriateness of different kinds of materials for various designs can be discussed with materials engineers, while properties such as tensile strength and fatigue life are documented in the engineering literature.

We can take specific steps to design for reliability. This is because failure generally is caused by inadequate design, manufacturing defects, use outside of specified conditions, or improper use. In each of these cases, the thoughtful designer can anticipate problems and "retire risk."

- *Inadequate design*, for example, can be addressed in part by material selection: is the material being used adequate for the intended use? In Chapter 12, we looked at the relationship between designing for strength and stiffness in our ladder. That material selection process can also inform us about the overall expected life of our design. While beyond our scope, professional engineers need to understand and consider material properties such as tensile strength, hardness, fatigue life, and creep in selecting materials that are used in critical environments.
- *Manufacturing defects* are best resolved by attention to DFM above, but we can also ask specific questions regarding the nature of proposed manufacturing and assembly processes and their impact on reliability. For example, are defects in manufacturing revealed during the process? If not, are there tests we can conduct to uncover them, or at least estimate their likelihood? When defects are revealed, are there procedures in place to allow us to reject substandard parts before putting them into use? If not, is development and application of such tests part of the design team's expected work?
- Looking at *use-related failures*, we can ask whether operation of our design is sufficiently clear to end users that the artifact will only be used in the manner and conditions intended. How do we know that? Has our testing and evaluation scheme given us insight into how the system might be used by someone outside the design team's orbit? Have we designed in a way that will permit visible labeling of dangerous or unsafe practices or environments?

Even at the design stage we need to consider the above questions if we are going to produce a reliable design. But that raises a final question—how reliable does the product need to be to meet our client's needs? Asking that question leads us all the way back to our consideration of objectives and metrics in Chapter 4. The proper metric for reliability is a number—the probability that the system will function under specified conditions for a specified amount of usage. The more critical the system is (e.g., air traffic control electronics), the higher that number will need to be.

### 14.2.2 Maintainability

Our understanding of reliability also suggests that many of the systems that we design will fail if they are used without being maintained, and that they may need some amount of repair even when they are properly maintained. This fact of life leads engineers to consider how to design things so that necessary maintenance can be performed effectively and efficiently. Maintainability can be defined as "the probability that a failed component or system will be restored or repaired to a specific condition within a period of time when maintenance is performed within prescribed procedures." As with our definition of reliability, we can learn from this definition.

First, maintainability depends upon a prior specification of the condition of the part or device, and on any maintenance or repair actions, which are part of the designer's responsibilities. Second, maintainability is concerned with the time needed to return a failed unit to service.

Designing for maintainability requires that the designer take an active role in setting goals for maintenance, such as times to repair, and in determining the specifications for maintenance and repair activities in order to realize these goals. This can take a number of forms, including:

- selecting parts that are easily accessed and repaired;
- providing redundancy so that systems can be operated while maintenance continues;
- specifying preventive or predictive maintenance procedures; and
- indicating the number and type of spare parts that should be held in inventories in order to reduce downtime when systems fail.

There are costs and consequences in each of these design choices. For example, a system may be designed with high levels of redundancy to limit downtime during maintenance, like an air traffic control system, but it may have very large attendant capital costs. Similarly, the cost of carrying inventories of spare parts can be quite high, especially if failures are rare. One strategy that has been increasingly adopted in many industries is to work toward making parts standard and components modular. Then spare parts inventories can be used more flexibly and efficiently, and components or subassemblies can be easily accessed and replaced. Any removed subassemblies can be repaired while the repaired system has been returned to service.

If high maintainability has been established as a significant design objective, design teams must take active steps in the design process to meet that goal. A design team should ask itself what maintenance actions reduce failures (especially in-service and catastrophic failures), what elements of the design support early detection of problems or failures (e.g., inspection), and what elements speed the return of failed items to use (e.g., repair). While no one would intentionally design systems to make maintenance more difficult, the world is fraught with examples in which it is difficult to believe otherwise, including a new car in which the owner had to remove the dashboard just to change a fuse.

Several considerations that designers should bear in mind when designing for maintenance and repair include:

- *Fault isolation and self-diagnosis*: It usually takes time to identify what has gone wrong with a system. As designers we can help reduce this time by building clear indicators into systems that identify the part of that system requires attention.
- *Part standardization and interchangeability*: Using standard parts in the design helps us to identify the number of parts held in inventories, and to reduce the skills needed to make repairs.
- *Modularization and accessibility*: Designs that *modularize* (i.e., package related components together) greatly reduce the time needed to restore a broken system to a working state, especially if the modules are themselves easy to replace. Parts with higher expected failure rates can often be placed in a system in ways that make them accessible without removal of other, functional parts.

If we follow these guidelines as members of a design team, we can better address our design problem, and we will gain a mindset that appreciates and respects users who must ultimately work with the system we are designing.

## 14.3 DESIGN FOR SUSTAINABILITY: WHAT ABOUT THE ENVIRONMENT?

Some people have come to hold negative views of technology and engineered systems because of the realization that one generation's progress may produce an environmental nightmare for the next. There are certainly enough examples of short-sighted projects (such as irrigation systems that created deserts or flood-control schemes that eliminated rivers entirely) that responsible engineers can feel at least some anxiety about what *their* best ideas might eventually produce. The engineering profession has come to appreciate these concerns over the past several decades, and has incorporated environmental responsibility directly into their codes of ethical obligations of engineers. The American Society of Civil Engineers (ASCE), for example, specifically directs engineers to "strive to comply with the principles of sustainable development"; the American Society of Mechanical Engineers (ASME) stipulates that "Engineers shall consider environmental impact in the performance of their duties." A number of tools to understand environmental effects are being introduced into engineering design in order to help with these issues and obligations (e.g., environmental *life-cycle assessments* (LCAs), which we discuss below). Some of these ethical obligations now take on the force of law, as we see in the requirement that many projects, especially public projects, conduct *environmental impact reviews* (EIRs).

### 14.3.1 Environmental Issues and Design

Environmental concerns relevant to design can be organized in any number of ways. Transportation engineering texts, for example, concern themselves with the impacts of engineered systems on water and air quality, while electrical engineering texts consider the effects of power generation and transmission, or focus on the particular environmental effects of some of the solvents and other chemicals associated with producing chips or printed circuit boards.

A more general approach is to think in terms of particular aspects of the environment and then consider the likely short- and long-term consequences of design alternatives.

We can often characterize the environmental implications of a design in terms of the effects on air quality, water quality, energy consumption, and waste generation. In each case, we need to address both short-term issues, which may arise as part of design for manufacturing or economic analysis, and long-term issues, which may not come up at all unless the designer raises them. Unfortunately, experience shows that the long-term effects of our design choices can completely overwhelm short-term benefits.

*Air quality* almost immediately springs to mind when we list environmental concerns related to design. Some urban areas have tremendous smog problems, small towns may have an industry with a large smokestack, and even national forests are experiencing loss of habitat due to acid rain and other air-quality issues. It is important to realize that these enormous problems often begin with relatively small emissions from various steps in the production of everyday objects. Each mile we drive in a car powered by a standard internal combustion engine adds a tiny bit of particulate matter, nitrous oxide, and carbon monoxide to the atmosphere. In addition, refining the fuel, smelting the steel, and curing the rubber for the tires add further emissions to the air. Less obvious but similar air quality problems result from the production of everyday materials in paper bags and plastic toys. In other words, designers concerned about the environment must consider both the manufacture of the product and its use.

Environmentally conscious engineers should also concern themselves with issues of *water quality* and *water consumption*. We may take the availability of clean water for granted, but many of the world's major bodies of water are under stress from overuse and pollution. As with air quality, this is a direct result of the multiple uses made of our water supplies. Many states have experienced severe droughts in the recent years, and in the southwestern United States, water is becoming the single biggest environmental constraint on further growth. Effective designers must consider and calculate the water requirements for producing and using their designs. Estimating changes to water resulting from particular designs is of great significance. These can include changes in water temperature (which for large-scale processes can affect fish and other parts of the ecosystem) and the addition of chemicals, particularly hazardous or long-lived compounds.

The production and use of designed systems requires *energy*. However, the energy demands of a system can be much higher than designers realize, or may come from sources that are particularly problematic environmentally. Several years ago, California faced an energy crisis that led to sporadic blackouts. Design choices about common household appliances such as refrigerators affect an increasingly energy-starved world. The variety of sizes, shapes, and levels of efficiency of refrigerators highlights the many design choices made by engineers and product design teams. Beneath the surface of such devices, however, there are further design choices made by engineers while generating and selecting alternatives. The principal energy consumer in a refrigerator is the compressor, which can be made more energy-efficient by judicious selection of components. Within the refrigerator walls, the use of insulation materials has a tremendous effect on how well cold temperatures are retained. Even door designs and their placement affects how much energy a refrigerator consumes. Designers must approach such projects systematically, applying all of the skills and techniques learned in their engineering science courses, and accounting for the consequences of their design choices.

Products must also be disposed of after fulfilling their useful life. In some cases, perfectly good designs become serious disposal problems. For example, consider the wooden railroad tie used to secure and stabilize trains tracks and distribute the loads into the underlying ballast. Properly maintained and supported, ties treated with creosote typically last more than 30 years, even under heavy loads and demanding weather conditions. Not surprisingly, most railroads use such ties. At the end of their lives, however, the same chemical treatment that made them last so long creates a major disposal problem. Improperly disposed of, the chemicals can leach into water supplies, making them harmful to living things. The ties also emit highly noxious, even toxic, fumes when burned. Thus, managing the *waste streams* associated with products and systems has become an important consideration in contemporary design. A great solution to one problem has become a problem in itself. The railroad industry has sponsored a number of research projects to explore ways to reuse, recycle, or at least better dispose of used ties, but the results remain to be seen. A more immediately obvious example of waste streams created by new technologies can be found in the many hand-held devices that most of us take for granted. Each new generation of cell phones, personal digital assistants, tablets, and personal computers also creates a new generation of electronic waste to be processed and disposed of. Sadly, there are environmental horror stories surrounding disposal of consumer electronics, particularly in developing countries, who often serve as the dumping ground for such products.

Sometimes the market fails to support the planned post-consumer disposal, even for products designed to be recyclable or reusable. Recycling is the intended final state of many paper and plastic products, for example, but many cities have found it difficult to successfully dispose of recycled paper and so are forced to place it in landfills. Battery companies have tried to develop recycling facilities to capture and control heavy metals and other dangerous waste products, but the small and omnipresent nature of batteries has made this very difficult.

### 14.3.2 Global Climate Change

Among the most pressing concerns facing us are the effects of climate change, sometimes referred to as global warming. There is overwhelming evidence that the average annual temperatures on the planet are rising, and a very strong consensus in the scientific community that human activity is responsible for some or all of this increase. The consequences of even modest increases in global temperature are likely to be catastrophic for some regions, such as polar ice caps, which are melting at surprising rates, and for some species that depend on particular climate conditions (such as polar bears). Engineers have a special obligation to involve themselves in finding ways to address global climate change, both because they have played a key role in the responsible technologies and because they have skills that can help moderate climate change.

One of the most important elements of global climate change is the extent to which carbon is emitted into the atmosphere, introducing what are referred to as "greenhouse gases." Many technologies emit carbon in ways that can surprise us, and when these technologies are used extensively the effects can be very significant. Some airplane engine designs, for example, emit very large volumes of carbon as a combustion by-product. Indeed, a pound of grapes flown from Chile to the United States results in six pounds of carbon being emitted into the atmosphere. Aircraft designers are working very hard to find ways to reduce carbon emissions from engines, but much work remains to be done.

Designing to reduce carbon emissions often begins with the measurement of the "carbon footprint" associated with producing the technology. The designer attempts to measure or estimate all the greenhouse gases emitted in all the processes to produce the product in question. This is still a very new analysis technique, and standards and methods are in flux at this time, but responsible engineers will be expected to understand and apply these techniques when designing for sustainability. As measuring the carbon footprint of technologies becomes better understood, the methods will find their way into life-cycle assessment, an important technique described in the Section 14.3.3.

### 14.3.3 Environmental Life-Cycle Assessments

*Life-cycle assessment* is a tool that was developed to help product designers understand, analyze, and document the full range of environmental effects of design, manufacturing, transport, sale, use, and disposal of products. Depending on the nature of the LCA and the product, such analysis begins with the acquisition and processing of raw materials (such as petroleum drilling and refining for plastic products, or foresting and processing of railroad ties), and continues until the product has been reused, recycled, or placed in a landfill. LCA has three essential steps:

- *Inventory analysis* lists all inputs (raw materials and energy) and outputs (products, wastes, and energy), as well as any intermediate outputs.
- *Impact analysis* lists all of the effects on the environment of each item identified in the inventory analysis, and quantifying or qualitatively describing the consequences (e.g., adverse health effects, impacts on ecosystems, or resource depletion).
- *Improvement analysis* lists, measures, and evaluates the needs and opportunities to address adverse effects found in the first two steps.

Obviously, one of the keys in LCA is the setting of assessment boundaries. Another is identifying appropriate measures and data sources for conducting the LCA. We cannot expect to find good, consistent data for all of the elements in the LCA, and so we must reconcile information from multiple sources. Because of differing project boundaries, data sources and reconciliation techniques, different analyses may produce different figures for the overall effects of a product, even when they're done in good faith. Therefore, it is very important that we list all the assumptions we've made and document all of the data sources we have used.

Currently, LCA is still in the early stages of development as a tool for engineering designers (and others concerned with the environmental effects of technologies). Notwithstanding its "youth," however, LCA is already a useful conceptual model for design, and is likely to become increasingly important for the evaluation of complex engineered systems.

## 14.4 NOTES

*Section 14.1*: This section draws heavily upon Pahl and Beitz (1996) and Ulrich and Eppinger (2000). In particular, the six-step process is a direct extension of a five-step approach in Ulrich and Eppinger (1995), with the iteration made explicit. Our discussion of DFA is adapted from Dixon and Poli (1995) and Ullman (1997); rules of assembly are cited in many places but are generally derived from Boothroyd and Dewhurst (1989).

*Section 14.2*: The definition of reliability comes from U.S. Military Standards Handbook 217B (MIL-STD-217B 1970) as quoted in Carter (1986). The failure discussion draws heavily from Little (1991). There are a number of formal treatments of reliability and the associated mathematics, including Ebeling (2010) and Lewis (1987). The definition of maintainability is from Ebeling (2010), as are the guidelines to design for reliability and maintainability.

*Section 14.3*: Codes of ethics for engineers are discussed further in Chapter 17. Sections 14.3.1 and 14.3.3 draw heavily on Rubin (2001), which also includes a very instructive example of LCA written by Cliff Davidson. The figure for carbon emissions from transporting grapes comes from McKibben (2007). Methodologies for calculating carbon footprints are given in Wiedmann and Minx (2007).

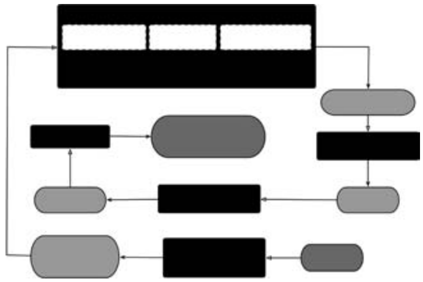
PART V

# *DESIGN TEAMS, TEAM MANAGEMENT, AND ETHICS IN DESIGN*

# CHAPTER 15

# *DESIGN TEAM DYNAMICS*

*We can do this together, as a team!*



WHILE INDIVIDUAL skills and personal responsibility are essential for success, the ability to work with others is a highly valued talent in successful design organizations. In this chapter, we examine how teams are transformed from a set of individuals into a group that can realize project goals. We also look at how teams can manage the conflict that almost inevitably arises when people work together.

## 15.1 FORMING DESIGN TEAMS

Design is an activity that is increasingly done by teams rather than by individuals acting alone. For example, product development teams include designers, manufacturing engineers, and marketing experts. These teams are assembled to bring together the diverse skills, experiences, and viewpoints needed to design, manufacture, and sell new products successfully. This dependence on teams is not surprising if we reflect on the design process (and its various stages) and design tools we have discussed in this book: Many of the activities and methods are devoted to achieving and acting on a common understanding of a design issue. Consider, for example, the difference between testing a structure in a laboratory and analyzing it with a computer model. While both activities require knowledge of structural mechanics, years of investment are required to master either the specific testing and laboratory skills or the right analysis and computer skills. We can achieve a lot when we assemble teams whose members bring their individual skills to work together successfully. We now briefly describe how teams form and perform, and how we can use this understanding to help with the design process.

### 15.1.1 Stages of Group Formation

Groups and teams are such an important element of human enterprise that they have been extensively studied and modeled. One of the most useful models of group formation suggests that groups undergo five stages of development that have been memorably named as:

- *forming*;
- *storming*;
- *norming*;
- *performing*; and
- *adjourning*.

We will use this five-stage model to describe some of the elements of group dynamics that are often encountered in engineering design projects.

***Forming***: Most of us experience a number of feelings simultaneously when we are initially assigned to a team or group. These feelings range from excitement and anticipation to anxiety and concern. We may worry about our ability—or that of our teammates—to perform the tasks asked of us. We may be concerned about who will show the leadership needed to accomplish the job. We may be so eager to get started that we rush into assignments and activities before we are really ready to begin. Each of these feelings and concerns are elements of the *forming* stage of group development, which has been characterized by a number of aspects and behaviors, such as:

- becoming oriented to the (design) task at hand;
- becoming acquainted with the other members of the team;
- testing group behaviors in an attempt to determine if there are common viewpoints and values;
- depending upon whoever is believed to be "in charge" of the project or task; and
- attempting to define some initial ground rules, usually by reference to explicitly stated or externally imposed rules.

In this stage, the team members may often do or say things that reflect their uncertainties and anxieties. It is important to recognize this because judgments made in the forming stage may not prove to be valid over the life of a project.

***Storming***: After the initial or forming stage, most groups come to understand that they will have to take an active role in defining the project and the tasks needed to complete it. At this point the team may resist or even resent the assignment, and it may challenge established roles and norms. This period of group development is known as the *storming* phase and is often marked by intense conflict as team members decide for themselves where the leadership and power of the team will lie, and what roles they must individually play. At the same time, the team will usually be redefining the project and tasks, and discussing opinions about the directions the team should explore. Some characteristics of the storming phase are:

- resistance to externally imposed task demands;
- interpersonal conflict;
- disagreement, often without apparent resolution; and
- a struggle for group leadership.

The storming phase is particularly important for the design team because there is often already a high level of uncertainty and ambiguity about client and user needs. Some team members may want to rush to solutions and will consider a more thoughtful exploration of the design space simply as a waste of time. At the same time, most design teams will not have as clear a leadership structure as, for example, a construction, manufacturing, or even a research project. It is important for effective teams to recognize when the team is spending too long in the storming phase and to encourage all team members to move to the next phases, norming and performing.

***Norming***: At some point, most groups do agree on ways of working together and on acceptable behaviors, or norms, for the group. This important period in the group's formation defines whether, for example, the group will insist that all members attend meetings, whether insulting or other disrespectful remarks will be tolerated, and whether or not team members will be held to high or low standards for acceptable work. It is particularly important that team members understand and agree to the outcome of this so-called *norming* phase because it may well determine both the tone and the quality of subsequent work. Some indicators of the norming phase include:

- roles in the group are clarified;
- informal leadership emerges;
- a consensus on group behaviors and norms develops; and
- a consensus on the group's activities and purpose emerges.

Significantly, norming is often the stage at which members decide just how seriously they are going to take the project. As such, it is important for team members who want a successful outcome to recognize that simply ignoring unacceptable behavior or poor work products will not be productive. For many teams, the norms of behavior that are established during the norming stage become the basis for the remainder of the project.

Many organizations use the *team charter* discussed in Chapter 16 to document or formalize the norms of the team, and to also articulate the overall scope and time scale of the project.

***Performing***: After the team has passed through the forming, storming, and norming stages, it reaches the stage of actively working on its project. This is the *performing* phase—the stage that most teams hope to reach. Here team members focus their energies on the tasks themselves, conduct themselves in accordance with the established norms of the group, and generate useful solutions to the problems they face. Attributes of the performing phase include:

- clearly understood roles and tasks;
- well-defined norms that support the overall goals of the project;

- sufficient interest and energy to accomplish tasks; and
- the development of solutions and results.

This is the stage of team development in which it becomes possible for the goals of the team to be fully realized.

***Adjourning***: The last phase that teams typically pass through is referred to as *adjourning*. This stage is reached when the group has accomplished its tasks and is preparing to disband. Depending on the extent to which the group has forged its own identity, this stage may be marked by members feeling regret that they will no longer be working together. Some team members may act out some of these concerns in ways that are not consistent with the group's prior norms. These feelings of regret typically emerge after teams (or any groups) have been working together for a very long time.

One final point about these stages of group formation should be made. Teams will typically pass through each of them *at least once*. If the team undertakes significant changes in composition or structure, such as a change in membership, or a change in team leadership, it is likely that the team will revisit the storming and norming phases again.

### 15.1.2 Team Dynamics and Design Process Activities

Our discussion of group formation helps us to understand how teams operate during the overall design process; it also provides insight into why some activities seem to succeed while others fail. Consider two activities that almost all design teams undertake: idea generation and report writing.

Teams that understand group dynamics are better equipped to choose good times to do team activities (e.g., the structured brainstorming we described in Chapter 2). In the forming and storming phases of team formation, members are likely still developing trust and confidence in each other. Indeed, the team as a whole may be trying to define the purpose of the project, while individual team members are worrying about their own roles in achieving that common purpose. The team may not yet agree on how seriously to take the project's initial goals, so it is likely not yet time to generate ideas as a group. On the other hand, during the norming stage, the team is developing a consensus about appropriate behavior, which makes respect-based techniques such as brainstorming possible. But perhaps the best time for a team to engage successfully in idea generation is during the performing stage.

Understanding team dynamics can also have a significant effect when the team is writing reports or preparing other documentation. In school and in practice, most engineers have had considerable experience in writing papers by themselves: We have all written term papers, technical memoranda, and lab reports. Documenting a design in a team setting is fundamentally different than writing a paper or lab report alone, because of our dependence on coauthors, on the technical demands, and on the need to ensure a uniform style. And, remember, final reports and other documents and presentations now reflect *on the team as a whole* and *on each team member.*

When writing as a team, we can only be sure what others are writing if all of our writing assignments and their associated content are explicit. *Every team member* should be involved in preparing outlines and rough drafts because they are essential to a team's success. Outline making in particular should be a team activity to ensure that every team member understands the overall flow of the paper, no matter how much each will write. It clarifies responsibilities and needs for cooperation that don't occur when individuals write alone.

Allocating work fairly doesn't guarantee the ultimate quality of a team's work products. Each team member should read draft reports carefully, and each team member must be allowed sufficient time to read and process draft materials at his own speed. No one on the team should be exempt from reading the final report drafts. The lead editor must listen to the individual needs of the team members, and they in turn need to speak up if they require time or any other resource to work on the report. This is clearly easier in an atmosphere of respect and trust. Equally important, the team must maintain an atmosphere in which the comments and suggestions of others are treated with respect and consideration. Given the pressure under which teams prepare final deliverables, the atmosphere tests the team's culture and attitudes. Teams need to closely monitor and carefully manage their interpersonal relationships.

Oral presentations also depend on successful team dynamics. Teams need to divide up work fairly, allowing every member to share in both the work and the glory. Team members should recognize that other members may be presenting their work. In some cases, the presenter of a particular piece of the project may have had little to do with the element of the work being presented, or may even have opposed that approach. Once again, then, the central issues here are: the need for every member to be familiar with *everything* the team has done; and team members must act appropriately and with mutual respect. Fortunately, final presentations come as projects end, by which time the team should have been in the performing stage for some time.

## 15.2 CONSTRUCTIVE CONFLICT: ENJOYING A GOOD FIGHT

Some degree of conflict is an inevitable by-product of people working together to accomplish tasks. Much of this conflict is healthy, a necessary part of exchanging ideas, comparing alternatives, and resolving differences of opinion. Conflict can, however, be unpleasant and unhealthy to a group, and it can result in some team members feeling shut out or unwanted by the rest of the group. A solid understanding of the ideas of constructive and destructive conflict is an essential starting point for team-based projects. Even in those cases where team members have been exposed to conflict management skills and tools, it is useful to review them at the start of every project.

The notion of constructive conflict had its origins in research on management conducted by Mary Parker Follett in the 1920s. She observed that the essential element underlying all conflict is a set of *differences*: differences of opinion, differences of interests, differences of underlying desires, and so on. Conflict is unavoidable in interpersonal settings, so it should be understood and used to increase the effectiveness of all of the people involved. To be useful, however, conflict must be constructive. *Constructive conflict* is usually based in the realm of ideas or values. On the other hand, *destructive conflict* is usually based on the personalities of the people involved. If we were to list situations where conflict is useful or healthy, we might find such items as "generating new ideas" or "exposing alternative viewpoints." A similar list of situations in which conflict reduces a team's effectiveness would likely include items such as "hurting feelings" or "reducing respect for others."

Teams ought to recognize and understand the difference between *destructive*, *personality-based conflict* and *constructive*, *idea-based conflict*. While a team is establishing norms—and even before these have been formalized or agreed to in the "norming" phase—it should establish some basic ground rules that prohibit destructive conflict and enforce them by responding if these rules are violated. Effective teams do not permit destructive conflict, including insults, personally denigrating remarks, or other such behaviors. If a team doesn't address this from the outset, it invites destructive conflict to become part of the team's culture.

Once we note this difference between constructive and destructive conflict, it is useful to recognize various ways that people react to and resolve conflict. Five basic strategies for resolving conflicts have been identified:

- *avoidance*: ignoring the conflict and hoping it will go away;
- *smoothing*: allowing the desires of the other party to win out in order to avoid the conflict;
- *forcing*: imposing a solution on the other party;
- *compromise*: attempting to meet the other party "halfway"; and
- *constructive engagement*: determining the underlying desire of all the parties and then seeking ways to realize them.

The first three of these strategies—avoidance, smoothing, and forcing—turn on the notion of somehow making the conflict "go away." Avoidance rarely works, and serves to undercut the other party's respect for the person who is hiding from the conflict. Smoothing may be appropriate for matters where one or both of the parties in conflict really don't care about the issue at hand, but it will not work if the dispute is over serious, important matters. Once again, the respect of the person "giving in" may become lost over time. Forcing is only likely to be effective if the power relationships are clear, such as in a "boss-subordinate" situation, and even then, the effects on morale and future participation may be very negative. Compromise, which is a first choice for many people, is actually a very risky strategy for teams and groups. At its core, it assumes that the dispute is over the "amount" or "degree" of something, rather than on a true underlying principle or difference. While this may work in cases such as labor rates or times allocated for something, it is not likely to be effective in matters such as choosing between two competing design alternatives. (We cannot, for example, compromise between a tunnel and bridge by building a suspension tunnel.) Even in those cases where compromise is possible, we should expect that the conflict is likely to reoccur after some period of time. Labor and management, for example, often compromise on wages, only to find themselves revisiting that same ground as soon as the next contract opens up. Thus, constructive conflict is the only idea that holds the possibility of stable solutions to important conflicts.

Constructive conflict takes the notion of identifying "the underlying desire" as its starting point. Each side must reflect on what it truly wants from a conflict and honestly report that to the other parties. Each side must also listen carefully to what the other party really wants. In many cases, the conflict is based not on the seeming problem, but rather on the fact that each party's underlying desires are different.

Follett told the following anecdote. She was working in a library at Harvard on a wintry day, with the windows closed. Another person came into the room and immediately opened one of the windows. This set the stage for a conflict—and for finding a way to resolve that conflict. Each of the five resolution alternatives outlined above was available, but most of them were unacceptable. Doing nothing or smoothing would have left Ms. Follett uncomfortably cold. Compromising by opening the window halfway did not appear to be a viable alternative. Instead, she chose to speak with the other person and express her desire to keep the window closed in order to avoid the chill and draft. The other party agreed that this was a good thing, but noted that the room was very stuffy, which in turn bothered his sinuses. Both agreed to look for a reasonable solution to their underlying desires. They were fortunate to find that an adjacent work area also had windows that could be opened, thus allowing fresh air to enter indirectly without creating a draft. Obviously, this solution was possible only because the configuration of the library allowed it. Nevertheless, they would not have even looked for this outcome except for their willingness to discuss their underlying desires. There are many cases where this will not work, such as when two people wish to each marry the same third person. There are, however, many cases where constructive engagement can succeed in enlarging the solution space available to the parties in conflict and heightening the understanding and respect of the other party. Even if a team is forced to revert to one of the "win–lose" strategies, it should always start by considering constructive engagement for resolving important conflicts.

## 15.3 LEADING DESIGN TEAMS

Most teams specify individuals as leaders at some point. Leadership roles can range from being the overall leader of the design effort to holding responsibility for a particular task, such as sanding the prototype in time for a major presentation. In this section, we look first at some key concepts associated with effective leadership. We will see that the concepts of team leadership and team membership are very closely tied together and can be understood using the notion of roles. We then look at some of the key roles that team members play in a successful team, and at how these roles are themselves leadership positions.

### 15.3.1 Leadership and Membership in Teams

Most of us have had the experience of working with both very effective and very ineffective leaders. Whether in a school setting, sports, a workplace, or some social activity, we encounter people acting as leaders. Our personal experience helps us to identify the behaviors exhibited by successful leaders and by unsuccessful leaders. We can draw several important lessons by thinking carefully about these behaviors and characteristics, and perhaps the most important of these is this: Effective leaders are not so much the result of having the "right" attributes as behaving in the "right" way. We will return to this point later in this section.

If we were asked to make a list of the behaviors and characteristics of *effective* leaders we have experienced, we might identify items such as being:

- open to other viewpoints;
- dedicated to the team's success;
- respectful of others;

- decisive, but only after considering all points of view;
- willing to do their share of the work; and
- knowledgeable about the team's subject.

Similarly, on a list of the behaviors and characteristics of *ineffective* leaders we have encountered, we might include:

- being closed-minded;
- bullying;
- expecting others to do more than the leader is doing;
- exhibiting interest only in their own field or topics; and
- not being respectful of others.

We could easily produce much longer lists of effective and ineffective leadership behaviors, but even these very short lists allow us to make some important observations. First, notice that what we hope to encounter in effective leaders is related to things they *do* rather that who or what they *are*. This suggests that effective leaders are made rather than born: We can each learn and practice actions and behaviors that make our team more effective.

Note too that the behaviors on our short lists, while recognizing technical knowledge, are not limited to technical skills. To use an analogy from sports, the highest scorer may not be best team captain. At the same time, understanding the scope and skills needed to complete the project cannot be overlooked. Successful team leaders need to have deep enough knowledge to understand and advance the project, and be able to respect the knowledge of others on the team.

Which brings us to a third point to note: Both lists have elements about listening to, respecting, and considering the viewpoints of others on the team. This suggests that we need to develop and practice "people skills" to be effective leaders. These "people skills" are often referred to as the *soft skills* of good engineers and designers, and they are much sought by industry.

Finally, it is instructive to notice that the behaviors of ineffective team *leaders* are not the behaviors we would want team *members* to exhibit. It is hard to imagine saying that "*A* is an excellent team member because he is a closed-minded bully." In fact, the behaviors we find in effective team leaders tend to be the same ones we find in effective team members. This suggests that some of the distinction between leadership and membership is artificial and related to team roles. If someone acts as the editor of the team's final report, is that person being a leader? The foregoing analysis would suggest that they are. One of the great opportunities provided by design projects in general, and student projects in particular, is that we can try out new roles and grow our leadership skills.

### 15.3.2 Personal Behavior and Roles in Team Settings

Because of the importance of teams in modern organizations, there are many formal models of roles and behaviors on teams. One approach that some engineering managers and educators have found useful is to focus on four behaviors that are critical to team success: communication, decision making, collaboration, and self-management. The

**TABLE 15.1 Critical behaviors and roles individuals can perform in teams**

| Behaviors | Roles |
|---|---|
| Communication | Active listener, influencer |
| Decision making | Analyzer, innovator, fact seeker |
| Collaboration | Conflict manager, team builder |
| Self-management | Goal director, process manager, consensus builder |

From McGourty and DeMeuse (2001).

importance of effective communication, for example, seems pretty obvious if we consider the need for each person to present to their teams and clients their problems, approaches, and solutions to the tasks for which they are responsible. Similarly, we need to be able to understand what other team members are working on and contribute to their activities. Nevertheless, most of us have personal experience with people who "just don't listen" or are so focused on their own contributions that they are planning their next remark while we are still speaking.

McGourty and DeMeuse have linked these critical behaviors to particular roles that individuals can play on the team, as shown in Table 15.1. Communication can be enhanced, for example, by taking on the role of active listener, in which we not only sit quietly while another person is speaking, but engage in actions that get them to clarify and explain their topic. Similarly, effective decision making requires people to analyze the situation, seek appropriate information to support the decision, and develop new ideas or alternatives. We have examined approaches to managing conflict earlier in this chapter, and we can see in this model a context for that discussion.

This approach to team behavior and roles raises an interesting opportunity for us as team members, even during the forming stage at the very beginning of a project. We can examine our own skills in light of the various roles and behaviors they support, determining what roles we like to play, which ones we are good at, and also which ones we want to improve upon. Going into a project with an open and honest personal appraisal makes us much likely to appreciate others, and gives us a framework to grow as a team member.

## 15.4 NOTES

*Section 15.1*: The model of group formation described here was initially developed by Tuckman (1965).

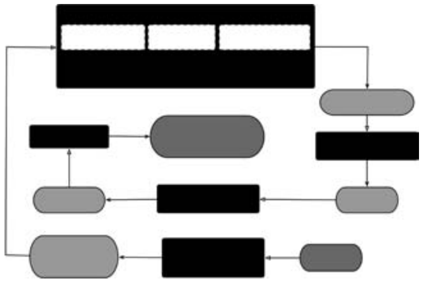
*Section 15.2*: Follett's essay introducing constructive engagement is reprinted in Graham (1996).

*Section 15.3*: Table 15.1 is taken from McGourty and DeMeuse (2001).

CHAPTER 16

# MANAGING A DESIGN PROJECT

*What do you want? When do you want it? How much are we going to spend?*



**DESIGN IS** an activity that can consume significant time and resources. In this chapter, we will present four tools that a small design team can use to manage and control a design project, particularly within an academic setting.

## 16.1 GETTING STARTED: ESTABLISHING THE MANAGERIAL NEEDS OF A PROJECT

All successful projects, whether large or small, whether on a grand scale or within an academic class, must address three essential concerns: The project team must (1) complete whatever tasks are required by the project, (2) do so within a specified time frame, and (3) work with the available resources. These three elements are summarized in terms of the "3S" model shown in Figure 16.1:

- The *scope* of a project sets limits on what the team must accomplish, particularly in prescribing the projects deliverables (e.g., a finished design, a working prototype, or fabrication specifications).
- *Scheduling* defines the time frame within which the project must be completed.
- *Spending* identifies and limits the available resources and how they may be applied to the project.

These project elements often compete with one another, so the project manager has to balance them in order to move the team to a successful outcome.

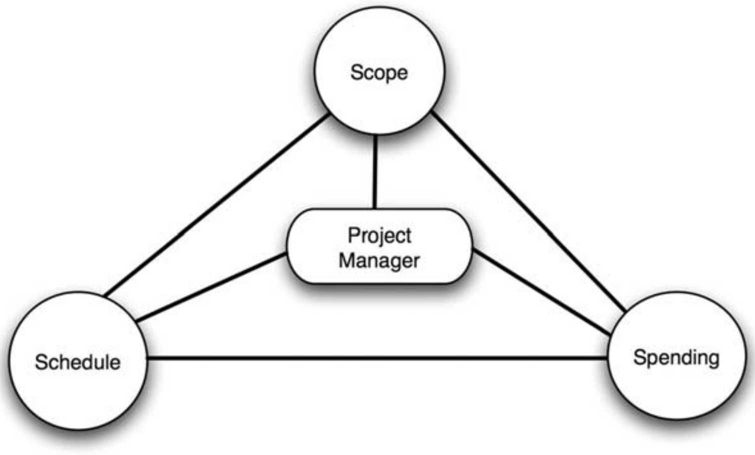



**FIGURE 16.1** Project management can be thought of as balancing the three Ss; Scope, Schedule, and Spending.

We noted in Chapter 2 that a design problem may begin with a simple client statement, "We want a design for a new product." The design team has to refine this open-ended statement into a clear list of *deliverables* consistent with the scope, schedule, and available resources. In direct analogy with the problem definition phase of the design process, successful project management begins with a questioning process aimed at articulating a clear understanding of what the client or sponsoring organization requires. Consistent with that analogy, we begin by asking questions—mostly, but not exclusively, of the client—in which aim to set the parameters for the project's three basic elements. So for our project we

*clarify the scope of the project* by asking questions like:

- Do you want a conceptual design, a detailed design, or something in between?
- Do you require a working prototype?
- Is full engineering functionality needed, or will a demonstration of key functions be sufficient?
- How much testing of the design and prototype do you want?
- Is a complete set of engineering drawings of the final design needed?
- How should the design be documented and formatted?
- To whom do we, the team, report?

*establish the schedule for the project* by asking questions like:

- When are the final deliverables due?
- Are there intermediate due dates?
- When should the team meet?

*identify spending and resource limits for the project* by asking questions like:

- Can we use outside resources (e.g., shop staff or drafting experts)?

- How many people will be on the team, and how much of their time is available for this project?
- What materials and supplies are available to the design team?
- Where will the team meet and work?

We can plan our team's work and avoid problems later if we get answers to these questions at the beginning of our project because we then can organize our time and resources to meet the project deliverables. Also, there will be no surprises at the end of the project about what was expected, when things were due, or how much the project cost. In Section 16.2 we look at two tools that are particularly helpful for clarifying and documenting the scope of a project.

## 16.2 TOOLS FOR MANAGING A PROJECT'S SCOPE

It would be nice if everything about a project could be known at the start. Then a project manager could simply list the work, add up the time duration, sum up the costs, and announce a work plan to the world. Unfortunately, most engineering projects don't work this way, and design projects face even more unknowns than most other projects: How long will it take to have a good idea? What will it cost to build a prototype for a design that is as yet unknown?

Put simply, design project managers face the complexities of other engineering projects, to which are added the uncertainties associated with creative activity. While this might tempt a design team leader to forego management tools, successful project managers have found that some specific tools can be very powerful in setting up and controlling projects. Two of these tools are (a) the team charter, which is a "mini-contract" between the team, its client, or sponsor, and the organization within which the design team works, and (b) the work breakdown structure, which is an ordered list of what must be done to complete all the required tasks to complete the project.

### 16.2.1 Team Charters

A team charter is an agreement that the team has reached with other stakeholders about the project requirements, including what constitutes project success, and what limits on the project may apply.

The team charter is, in many ways, the management analog to the revised problem statement that we discussed in Chapter 3. Recall that at the outset of the project there may be errors, biases, and incomplete information in the initial problem statement. An effective design team works to restate the problem in a way that clearly states the client's needs while offering a large design space within which to meet those needs. In a similar vein, a project manager needs to understand:

- the goals of the project, including both minimally acceptable goals and "stretch" goals;
- how those goals align with larger organizational goals;
- the authority for the project;

- the project deliverables;
- the time frame for the project, including any limits on the schedule;
- the resources available for the project; and
- any unusual circumstances associated with the project.

This information, once gathered, can be written up in the form of a team charter and distributed to the client, managers within the larger organization, and all team members. In many organizations, drafting of the charter is begun before the team (or even a project manager) is selected, and is then completed by the project manager. Often the sponsor, senior managers of the design firm, and the project manager sign the charter or agreement to complete the chartering process.

The charter is used for a number of purposes throughout the project's life. It describes the project to prospective team members, elicits commitment from team members, and settles conflicts about resources, timing, or scope. Because of its written nature, it can also be used later in the project to avoid "scope creep" or "mission growth," the tendency for project goals to be expanded as the problem becomes better understood, or as team or client interests change.

Like the revised problem statement, team charters are fairly short, usually only one or two pages long. It needs only to be long enough to lay out the goals, roles, resources, deliverables, and major schedule points. The key is that the project manager and the team include the information needed to make the charter valuable as a negotiating tool and a formal record when it may be needed later.

Figure 16.2 shows an example of a team charter for a student project. Notice the charter's overall structure. It first lays out the goals for the project in terms of the client or sponsor's goals, the team's goals, and the organizational goals (in this case, the course instructors). This helps identify any disconnects between the parties and also clarifies what each party hopes to get out of the team's work. The charter then specifically spells out the deliverables, or work products, for which the team is responsible. This sets limits on what must be done and helps the project manager define the resources needed to complete the project. Many charters also incorporate a schedule of key milestones in the deliverables section by specifying due dates. The charter also calls out the resource limits for the project. This is important so the client does not misunderstand the intended work effort—thinking, for example, that the student team is working full time on the project. It also commits the team members to full participation at the level of a course, avoiding free rider problems. Many charters will also include a section for project-specific remarks that may be important to some of the parties. Finally, the charter, while not truly contractual in nature, may include a signature block to help commit the team and other stakeholders to the project's intentions.

The information used to develop the charter is usually readily available to project managers early in the project. For example, the aforementioned revised problem statement should identify project goals and deliverables, and it may include schedule constraints. Wise project managers also know to negotiate resources early in the process of taking on a project; the project manager's authority and any special circumstances are also likely to emerge during these early discussions.

It is important to note that project goals and organizational goals are distinct objects: *Project goals* are the specific accomplishments or outcomes expected of the project itself,

**DANBURY SCHOOL ARM SUPPORT PROJECT**
**TEAM CHARTER**

This charter documents key information regarding a project to be conducted by students at Harvey Mudd College (HMC) on behalf of a student at the Danbury School. The faculty advisor will be Professor Dym; the liaisons for Danbury will be Mssrs. H. and J. The project manager/team leader will be selected at the outset of the project by and from among the members of the student team.

The team agrees to abide by all restrictions and regulations of Danbury School when on site, to place the safety of the client as its highest priority, and to work in accord with HMC's Honor Code.

**Goals**

The project is assigned as part of HMC's introductory design course, and the students understand that they are expected to work to accomplish HMC's goals for the course, Danbury's goals for the project, and the needs of a particular user.

The goals of the E4 course are to:

1. develop an understanding and experience of the conceptual design process;
2. give the students experience in team dynamics; and
3. enable students to learn to manage small design projects.

The goals of the team are to:

1. satisfy the client' s needs and meet E4 course requirements;
2. learn about engineering and engineering design; and
3. have fun while doing a good thing for someone.

The goals of the Danbury school are to:

1. improve Jessica' s learning environment and quality of life; and
2. increase public awareness about the conditions faced by students like Jessica.

**Deliverables**

The following deliverables will be completed by the last Friday before final exam week:

1. a prototype arm stabilization device which has been designed, built and tested for Jessica' s use;
2. final design documentation for Danbury School and for the E4 teaching team; and
3. a public presentation of the team' s design process and results.

**Resource Limits**

The team is expected to work an average of 10 hours per week per team member. HMC will provide not more than $125 to the team for the purpose of purchasing supplies and materials for the project.

**Other Restrictions or Information**

The team will place the safety of the user above all other aspects in their design activity. The team will hold weekly meetings with their faculty advisor, and will meet regularly with the sponsors at Danbury.

**FIGURE 16.2** A team charter for the Danbury support arm project sets out goals, relevant parties, deliverables, and time and resource limits.

such as a solution to a specific problem. *Organizational goals* reflect the intentions of the larger entity (such as the college, for a student project). These organizational goals may include exploring potential technologies or developing skills that will be useful in future endeavors. If project goals and organizational goals are not aligned, it is quite possible that a project may seem like a success but ultimately go nowhere—because the organization chose not to pursue it. If the project and organizational goals are aligned, it may be that a project that appears relatively unexciting on its own will be seen as part of a larger, more exciting program.

### 16.2.2 Work Breakdown Structures

Most engineers address complex problems by breaking them into smaller problems, with decomposition continuing until each piece can be solved with known methods, or until the hardest part of the problem has been characterized. In a similar fashion, engineering managers like to decompose projects into small enough units that they can identify who will be responsible, establish how long the smaller task will take, and determine the resources required to complete that smaller task (or subtask).

The primary tool that is used to determine the scope of a project is the *work breakdown structure* (WBS). The WBS is a hierarchical representation of all of the tasks that must be performed to complete a design project. Project managers generally use WBSs to break the work down (hence the name) into pieces sufficiently small that the resources and time needed for each task can be estimated with confidence.

Consider the everyday task of driving a car. Even though this is a common task, many of us might be overwhelmed if asked to describe *exactly* how to start and drive a car. We might start with something like "first you adjust the seat and mirrors"—which presupposes that you have already gotten into the car and thus know how to get in. If forced to describe this task to someone who rarely drove or relied exclusively on public transportation, we might break our description down into a number of task groups, such as unlocking the car door, getting into the car, adjusting the seat and mirrors, starting the engine, maneuvering the car, and so on. We might even want to review the entire plan before starting the car so that our novice driver is comfortable with all of the steps between unlocking the car at the start and safely exiting at the finish. This decomposition of tasks and subtasks is the central idea of the work breakdown structure. When we are confronted with a new, large, or difficult task, one of the best ways to figure out a plan of attack is to break it down into smaller, more manageable subtasks.

A more compelling example is that of a team that has been asked to design a spacecraft. The team will have to design across several specialties, including propulsion, communications, instrumentation, and structures. Here the design team leader will want to be certain that all the necessary tasks have been listed and included to ensure that there are no surprises when the craft is in flight. The WBS lists all of the tasks needed to complete the project and organizes them so as to help the project leader and the design team understand how all those tasks fit into the overall design project.

Figure 16.3 shows a work breakdown structure for the tasks from the juice container design. At the top level, the WBS is organized in terms of eight basic task areas:

- understand customer requirements;
- analyze function requirements;

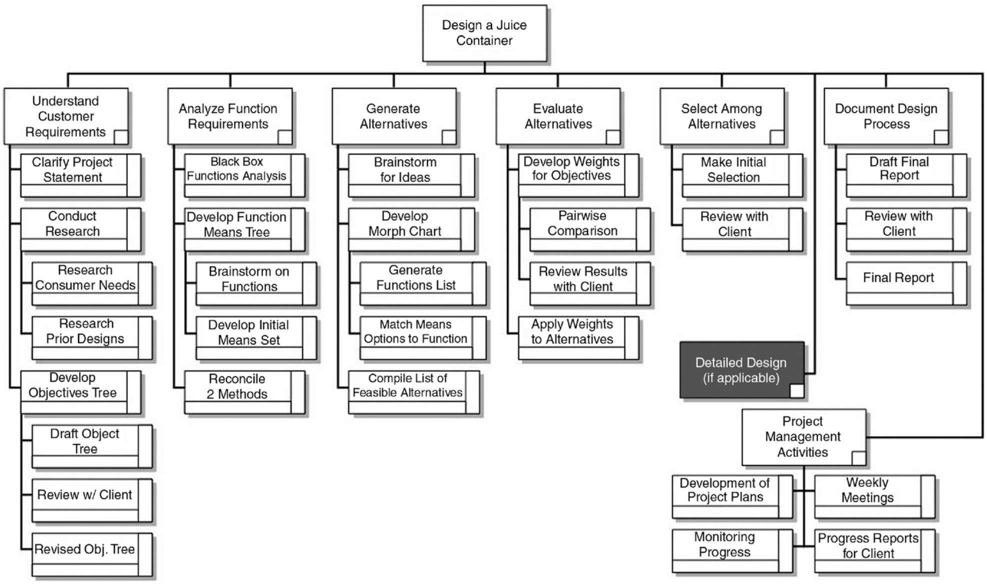



**FIGURE 16.3** A WBS for the juice container design project. Because the design project is just beginning, the structure necessarily takes on a formal and somewhat generic framework. Note, however, that the designers are already aware of some details, such as the distinction between identifying consumer needs and prior designs.

- generate alternatives;
- evaluate alternatives;
- select among alternatives;
- document the design process;
- manage the project; and
- detailed design.

Each of these top-level tasks has been broken down in subtasks. Because of page size limitations, we show in great detail only a few of the tasks (e.g., understanding customer requirements) for this example. If we were actually on a team carrying out this project, we would likely go into greater depth in all of the areas. We look for four properties in a WBS such as that displayed in Figure 16.3:

1. *Full breakdown into parts*: Each item that is taken to a lower level is always broken down into two or more subtasks at that lower level.
2. *Adequacy*: The breakdown is adequate to allow the parts to be managed. If we cannot determine how long an activity will take or who will do that activity, then a key WBS rule is that we should break the task down further. In fact, experienced project managers will be more inclined to have shorter, less detailed WBSs than relatively inexperienced managers, because their greater experience makes them more able to aggregate subtasks into identifiable, measurable tasks.
3. *Completeness*: Any task or activity that consumes resources or takes time should be included in the WBS explicitly or as a known component of another task. That is why the tasks of documentation and management are shown in Figure 16.3. Activities such as writing reports, attending meetings, and presenting results are essential to the completion of the project, and failure to plan for them as work will certainly result in problems later.
4. *Additive*: All the lower subtasks of the hierarchy must be equivalent to completion of the full task above them. This also means that the time needed to complete an activity at a top level should be the sum of the times for tasks listed at the level below. Therefore, breaking work down to the next level below must be done thoroughly and completely.

It is also important to note what the WBS is *not*. First, a WBS is *not* an organization chart for completing a project. Although "org charts" are visually similar, a WBS is a breakdown of *tasks*, *not* of titles, roles, or people in an organization. Second, a WBS is *not* a flowchart showing temporal or logical relationships among tasks. The listing of tasks in a WBS may be organized such that a task (e.g., writing the final report) is shown in a different part of the hierarchy than other tasks that must precede it (e.g., all of the design, building, and testing that is being reported). Third and last, a WBS is *not* a complete listing of all the disciplines or skills required to complete the project. Tasks performed by team members with many different skills can be combined into the same part of the hierarchy as long as the WBS meets the above criteria.

Figure 16.4 shows a sample WBS for the Danbury arm support project. It is organized as an indented list rather than a graph, and the tasks are arranged differently than the juice container example, with areas like "Preliminary work," "Zoning in," and

**Work Breakdown Structure**

I. Preliminary Work
  a. Examine problem statement
  b. Research
  c. Visit Danbury School
    i. Talk with Jessica
    ii. Take measurements
    iii. Hands-on examination of what device must do
    iv. Talk with teachers and physical therapist
II. Zoning In
  a. Finalize problem statement
  b. Create full list of objectives and constraints
  c. Rank objectives
  d. Create metrics
  e. Create objectives tree and determine functions
  f. Communicate with client on objectives, constraints, and functions
III. Brainstorming for Design Alternatives
  a. Use morph chart to create combinations of alternatives
  b. Create sketches of design alternatives
  c. Conceptual testing of alternatives
IV. Picking a Design
  a. Use metrics and objectives tree
  b. Compare the design alternatives against metrics
  c. Communicate with client
  d. Preliminary testing on selected design
V. Building a Prototype
  a. Activities prior to building
  b. Gather materials
  c. Determine sources and tools needed to build
  d. Divide work
    i. Building prototype
    ii. Assemble parts
    iii. Test prototype
    iv. Document test results
VI. Final Report
VII. Presentations to client and class
VIII. Overarching Work
  a. Organize meetings
  b. Organize building days

**FIGURE 16.4** An excerpt from a WBS for a third Danbury support arm team. Notice that the overall structure is not the same as that in Figure 16.3, but many of the same tasks appear. Ahmad et al. (2007).

"Overarching work." The key point here is that a team can arrange the WBS in some other way (i.e., the WBS does not have to be a graph), so long as it reflects the four basic properties: full breakdown into parts, adequacy, completeness, and additive. In fact, teams often find it easier to start with an indented outline when they first draft a WBS. We will see later, however, that the graphical form has some advantages for monitoring a project's progress.

In the end, the WBS is a tool for a project team to use to ensure that they understand all of the tasks that are needed to complete their project. That is why a WBS is so valuable for determining the scope of the project.

## 16.3 THE TEAM CALENDAR: A TOOL FOR MANAGING A PROJECT'S SCHEDULE

Scheduling tools help us to plan activities, remind us of important project due dates, and also assist in identifying those things that might really mess up our project if we don't do them on time. We focus here on the team calendar, which shows the time that is available to the design team, with highlights that indicate deadlines and time frames within which work must be completed. Team calendars are used by virtually all teams, particularly those doing design work, and it is the most suitable scheduling tool for doing conceptual design projects in an academic setting.

A *team calendar* is simply a mapping of deadlines onto a conventional desk or wall calendar. Such deadlines will certainly include externally imposed ones, such as commitments to clients (or to professors for academic projects), but should also include team-generated deadlines for the tasks developed in the WBS. In this sense, the team calendar is really an agreement by the team to assign the resources and time necessary to meet the deadlines shown on the calendar. Figure 16.5 shows a team calendar for a student design team that is seeking to complete its project by the end of April, an externally imposed deadline. Note that the calendar includes several deadlines over which the team probably has no control, such as when the final report is due and when in-class presentation of results is to be done. It also includes routine or recurring activities, such as Tuesday night team meetings. Finally, it includes some deadlines that the team has committed to realizing, such as completing a prototype by 5 p.m. on April 2.

Several points should be kept in mind in setting up a team calendar. First, the idea of a team calendar implies that the deadlines are all understood and agreed to by everyone on the team. As such, the calendar becomes a document that can—and should—be reviewed at every team meeting. Second, the team calendar should allow times that are consistent with the time estimates generated in the WBS. If a task was determined to take two weeks to complete, there is little point in allowing that task only one week on the team calendar. A final point to note is that a team calendar, while easily understood by members of a team, *cannot by itself capture the relationship between activities*. For example, in Figure 16.5 we see that building the prototype precedes proof-of-concept testing *only* because the team chose to put it that way. For many devices a proof-of-concept may precede building a final prototype. A team calendar cannot address this sort of problem, nor can it "remember" team decisions of this sort.

It is important to note that the WBS and the team calendar are related. Although the WBS is not structured to reveal the temporal relationships among tasks, it can often remind us of them, especially if we have followed the principles of completeness and adequacy. For example, just noting that work on a foundation includes digging a hole, erecting forms, and pouring concrete, does not guarantee that we will do this in order, but it certainly should help us remember. Similarly, the adequacy property says we should have broken our tasks into pieces for which we can assign responsibility and estimate durations. This means

**Design Team**

March

| S | M | T | W | T | F | S |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| 21 | 22 | 23 | 24 | 25 | 26 | 27 |
| 28 | 29 | 30 | 31 | | | |

May

| S | M | T | W | T | F | S |
|---|---|---|---|---|---|---|
| | | | | | | 1 |
| 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 9 | 10 | 11 | 12 | 3 | 14 | 15 |
| 16 | 17 | 18 | 19 | 20 | 21 | 22 |
| 23 | 24 | 25 | 26 | 27 | 28 | 29 |
| 30 | 31 | | | | | |

**April**

| Sun | Mon | Tue | Wed | Thu | Fri | Sat |
|---|---|---|---|---|---|---|
| | | | | 1 | 2<br>**5:00PM Prototype Built** | 3 |
| 4 | 5 | 6<br>**7:00–8:15PM Team Meeting** | 7 | 8 | 9<br>**11:00AM Proof of Concept Due** | 10 |
| 11 | 12<br>**11:00AM Rough Outline Due** | 13<br>**7:00–8:15PM Team Meeting** | 14 | 15 | 16<br>**5:00PM Topic Stce Outline Due** | 17 |
| 18 | 19<br>**11:00AM Prsntion Outline Due** | 20<br>**7:00–8:15PM Team Meeting** | 21<br>**11:00AM Slides Due** | 22 | 23<br>**5:00PM Draft Final Report Due** | 24 |
| 25 | 26<br>**10:00–11:00AM Present Results** | 27<br>**7:00–8:15PM Team Meeting** | 28 | 29 | 30<br>**5:00PM Final Report Due** | |

**FIGURE 16.5** A team calendar for a student design project that includes externally imposed deadlines, team commitments, and recurring meetings. It is usually better to make the team calendar "too complete" than it is to leave out potentially important milestones or deadlines.

a project manager can, while building a team calendar, go to the responsible individual to ask about timing. In order to facilitate this, many teams modify their WBS to include a box for each task and subtask showing the duration. Using the additive property, we should only need durations for each of the lowest level tasks.

Finally, we note two other scheduling tools that are frequently used in project management: the *Gantt chart*, which is a horizontal bar graph mapping activities against a time line; and the *activity network*, which graphs the activities and events of the project, showing the logical ordering in which they must be performed. Gantt charts and activity networks are both powerful graphical representations of the logical relationships between tasks and the time frames in which they are to be done. Both are used for large-scale projects and are almost always generated using project management software.

## 16.4 THE BUDGET: A TOOL FOR MANAGING A PROJECT'S SPENDING

A *budget* is a list of all of the items that will incur an economic cost, organized into some set of logically related categories (e.g., labor and materials). The budget is the key tool for managing spending activities in a project. Note that there is an important distinction between the budget for doing the design, or design activities, and the budget needed to produce or build the artifact being designed. Our concern is primarily with the budget needed to do a design.

Budgets are difficult but essential tools for project management. They permit teams to identify the financial and other resources required, and to match those requirements to the available resources. Budgets also permit teams to account for how they are spending project resources, some of which are reflected in money, others in hours (e.g., student time on an academic project). Finally, budgets also formalize the support of the larger organization from which a team is drawn.

Fortunately, we usually don't need large, complex budgets for doing the sort of design projects we are likely to do in an academic setting. Remember that we are concerned with the *budget for doing the designs*, not with the budget for making the designed objects. Figure 16.6 shows what a budget for the Danbury School project might look like. Design project costs normally include labor, research expenses, materials for prototypes, travel, and other support expenses. We limit our budget discussions to these cost categories, looking specifically at labor, materials, travel, and incidental expenses.

Estimating labor budgets usually begins with analysis of each type of labor required for the tasks of the project, estimated in hours of work, and only then multiplied by a labor rate. For student projects, it makes more sense to estimate the hours involved than worry about wage rates, since students generally don't get paid to do their coursework. A good starting point for estimating labor hours are the durations implied in the WBS, although this is not sufficient. Because the WBS provides durations, no distinction is drawn between a one-week task done by the whole team and a one-week task done by one team member. This means that the WBS must be broken down into more detail to identify the number of people working on a task during its duration as specified in the WBS.

In budgeting for materials in a design project, it is necessary at the outset to think about what sorts of solutions are *possible* and also to keep in mind what the charter requires. If, for example, a functional prototype is a deliverable, we need to plan and

**Danbury School Project**
**Proposed Time and Materials Budget**

| | |
|---|---|
| Student Labor: | |
| 10 hours/wk • 6wks • 4 students | 240 Hours |
| @ $0/hr | $0.00 |
| Travel: | |
| 6 trips to Danbury • 3miles • $0.50/mile | $9.00 |
| Materials Budget: | |
| Not to exceed $125 (less travel) | $116.00 |
| Total Budget: | |
| Not to exceed $125 | $125.00 |

**FIGURE 16.6** A team budget for the Danbury support arm project. The budget allocates hours to the project for labor, and uses dollars to track out-of-pocket expenditures. In a company, the hours would typically be translated into dollar costs. Notice that the team has been given a not-to-exceed amount, so it is allowed only $116 for materials for its prototype.

estimate costs to build and deliver it, including labor and materials. Similarly, if the final report is to be bound in a particular way, we should budget for that cost. This is not to say that we have implied a particular design solution; rather, we should consider our resource needs as early as we can. To address this uncertainty, design teams often use *not-to-exceed* budgets for design projects. This type of budget specifies an upper limit on what costs are allowed. There are two dangers in this approach: First, if an organization budgets this way routinely, resources set aside for design projects may never be used. Second, teams may overspend, viewing not-to-exceed as "Why not? We have budgeted this much money."

Travel expenses arise if we need to meet with clients or users, or if we have project needs that can only be met with off-site visits. Most organizations have internal rules about travel (e.g., mileage allowances, per diem limits on meal and lodging). Thus, it is a good idea to check with internal supervisors (e.g., faculty advisors) before making travel budgets.

Many budgets include a "miscellaneous" category. This reflects a temptation to set aside resources, time, or money, to balance later overruns in other project categories. We should avoid this trap. Using such "fudge-factors" muddies our understanding of project cost structure, needlessly ties up organizational resources, and reflects a lack of clear thinking on our (and our project manager's) part. We should only call out specific costs we expect to incur (e.g., library-related research expenses). Using safety factors or a "miscellaneous" category to hide reserves reflects poorly on our team and our management skills.

Finally, it is important that we properly value the time invested in a design project by each and every member of a design team. This is important even in projects done in design courses. (In fact, there is a tendency to undervalue this very scarce resource simply because we typically do not attach a dollar figure to student labor in the budget.) In practice, most engineering firms charge between *two and four times* an employee's direct compensation when they bill a client for that employee's time. That multiplier covers fringe benefits, overhead costs, supervision, and profit. If we were to pay student time at a wage rate of only

\$8.00 per hour, a team of four students working 10 h each week on a project for 10 weeks would be billed out by a design firm at \$6400–\$12,800 for the entire project. Put in dollar terms, time is a valuable, scarce, and irreplaceable resource—don't waste it! Put in human terms, when we undervalue the time of our fellow team members, we are disrespecting the contributions we expect them to make to our project's success.

## 16.5 MONITORING AND CONTROLLING PROJECTS: MEASURING A PROJECT'S PROGRESS

We have now developed a plan, expressed in terms of a charter, a schedule, and a budget. How do we track our team's performance relative to that plan? This is an important question, but it can be very hard to answer even for a simple academic design project. If we were managing a construction project, we can go out and see if a task has been done by the date planned. In a design project, however, monitoring and control are more subtle and, in some ways, more difficult. Therefore, it is essential that our team agree on a process for monitoring our progress before the project gets very far underway.

There are a number of techniques and tools available to monitor projects, but these often involve team members filling out time sheets, punching clocks, or other accounting tools. For smaller design projects, especially academic projects, these kinds of tools may not be very effective. We therefore present a highly simplified version of the *percent-complete matrix* (PCM) that is widely used in an industry to relate the extent of work done on the parts of a project to the status of the project overall.

The goal of the PCM is to use the information in the WBS and, if available, the labor budget, to determine the overall status of the project. Constructing a PCM requires only that we know the cost (or time requirements) of each item or area of interest, and the percent of the total cost or time corresponding to that item. The PCM then allows us to input the percent of the work on that task or work item and, by summing over all the items in the project, we can calculate the total percent of the project completed. In general, the method is best suited to cases where there some clear method of calculating progress is available. If, for example, the foundation work of a building project constitutes 25% of the total expected costs of a project, then completing one-half of the foundation work means we have completed 12.5% of the cost basis of the project. A manager can periodically update progress in each of the general areas in the WBS to determine overall project progress.

In design projects a physical measure such as yards of concrete poured is not available for estimating progress on a task. In fact, design managers are generally more concerned with progress relative to allowed time than to available budget or any physical measures. One way around this is to use the tasks in our WBS as a control tool. Recall that we estimated the time duration of each activity, and that completion of all the lowest level activities of the WBS sums to completion of the project. This means that if we track progress along the bottom of the WBS, we will have tracked progress for the entire project.

Some design teams use a simple algorithm for tracking task-level progress. When work is begun on a task, a share of that task's hours (e.g., 25%) is immediately credited to the team. At team meetings, the responsible individual for that task is asked for updates. When she says that the work is half finished, the task is credited with 50% progress. When the responsible individual says the task is finished, 90% is credited, and the remainder is credited to the task when the relevant work product is accepted by the team or the project manager.

Consider the simplified WBS in Figures 16.7 and 16.8. Figure 16.7 shows a total time estimate for each task. Figure 16.8 shows the progress on each task with a graphic legend, as follows: Tasks are highlighted in bold when begun, marked with a slash when completed halfway, marked with an X when turned in, and shaded when accepted. If we use this crediting scheme, we can estimate the percent of the project that has been completed. In this case, the team has held meetings with its client (32 h), completed its research (120 h), held half of its meetings with users ($32 \times 50\% = 16$ h), and half of developing its objectives

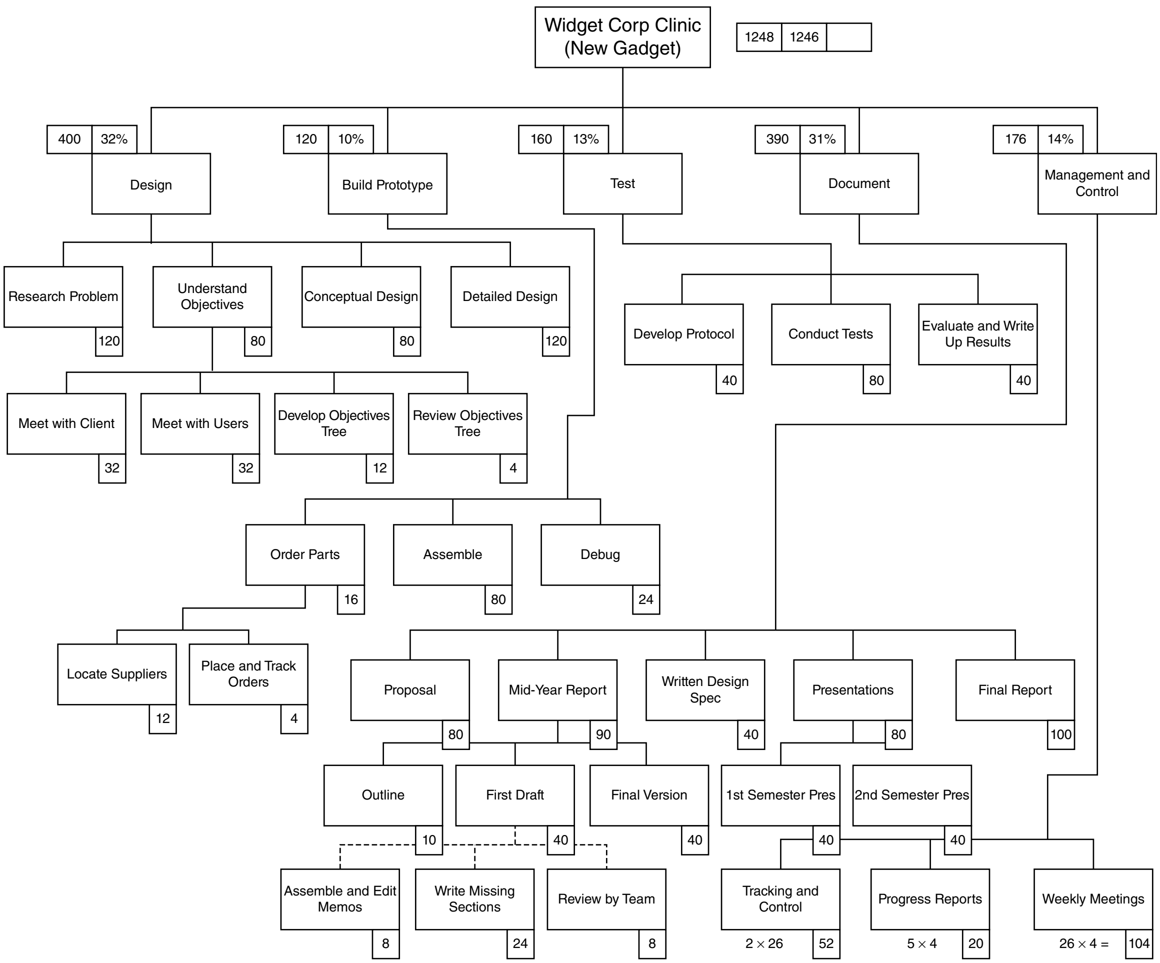



**FIGURE 16.7** A WBS with allocations for tracking team progress in hours. Notice that the hours for the lower subtasks sum up to the hours for the level above, which is consistent with the requirements for WBS. With this version of the WBS, the team can use cross-outs to indicate an estimate of the percent of work completed.

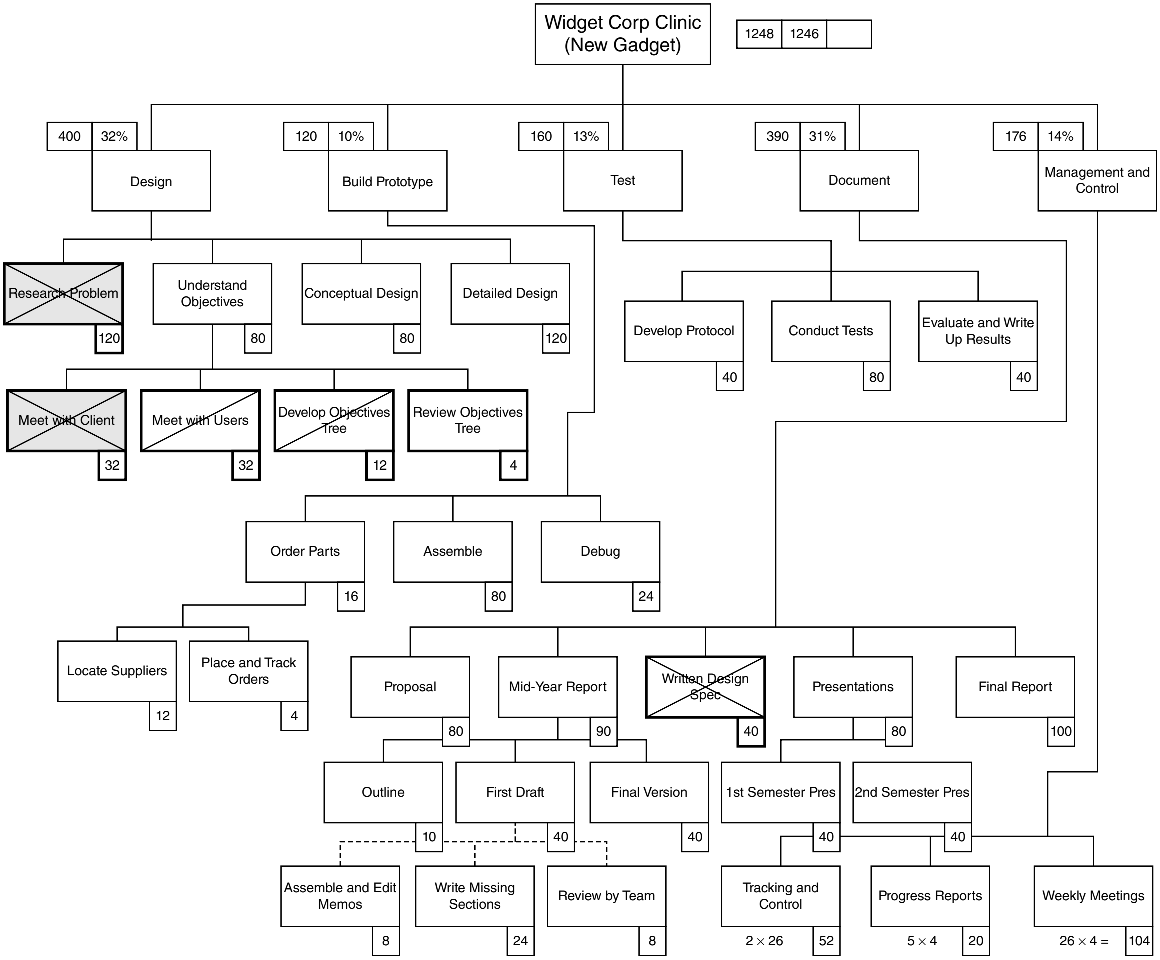



**FIGURE 16.8** The WBS in Figure 16.7 used as a monitoring tool. The boxes in bold represent tasks which have begun (credited with 25% of the task hours), those with slashes are half finished (50%), those with Xs have been turned in as complete by the responsible individual (90%), and those shaded have been accepted by the project manager or client (100%).

tree ($12 \times 50\% = 6$ h). They have also begun reviewing drafts of the objectives tree ($4 \times 25\% = 1$ h) and have completed the written design specifications task—although the final product has not yet been accepted ($40 \times 90\% = 36$ h). Adding up the numbers, the team has completed 211 h ($32 + 120 + 16 + 6 + 1 + 36$). Since the total project time is 1248 h, the team has completed 17% (211/1248) of the project. This may represent good progress if we are monitoring early in the project. If we're measuring halfway through the project's projected calendar, the results are likely not so good.

This simple approach allows a project manager to monitor progress in several ways. By adding up (percentage × duration) products, the manager can estimate the overall task share done and compare it to the actual calendar time used. The manager can also note whether there's a problem with a particular team member. If, for example, A always begins projects (bolded boxes) but never completes them (X's), the manager (and the entire team) will see this as the project unfolds.

We can't, however, monitor design projects just with charts. We need team meetings, private conversations, and techniques such as those we presented in Chapter 15, to motivate our team. At times, we may even use creative conflict. Nevertheless, formal tools are useful to teams, both to describe progress and to serve as a basis for conversation in the person-oriented aspects of management.

## 16.6 MANAGING THE END OF A PROJECT

In practice, most projects do not simply end with the delivery of materials to the client, but with a *project postaudit*: An organized review of the project is conducted, including its technical work, management practices, work-load and assignments, and its final outcomes. This is an excellent practice to follow, even for student projects that end with teams disbanding completely.

The key point behind a postproject audit is to focus on doing an even better job next time around: What mistakes did we make that we can avoid in the future? Are there particular things we can do better? As a practical matter, a postproject audit may happen in simple hour-long meeting, or it may be part of a larger formal process that is directed by a design team's organization. Regardless of the mechanism, there are four elements in a basic postaudit process:

- review the project goals;
- review the project processes, especially in terms of ordering of events;
- review the project plans, budgets, and use of resources; and
- review the outcomes.

Reviewing the project goals is particularly important for design projects, since design is a goal-oriented activity. If the project was supposed to solve problem A, then even an idea that results in earning a patent for solving problem B may not always be viewed as a success. We can only evaluate a project meaningfully in terms of what it set out to do. To this end, we should review how we used the problem definition tools and techniques as part of our postaudit.

Closely tied to reviewing the results of using design (and management) tools is evaluating the effectiveness of the tools themselves. Just as a toolbox may contain many items that are only useful some of the time, many of the formal methods and techniques that are presented in this book and elsewhere will be more effective in some situations than in others. No catalog of successes or failures by outside authors will have the same purchase with a team as its own experience. Reflecting on what worked and what didn't and coming to grips with why a tool did or did not work are important elements of a postproject audit.

Similarly, it is also important to review how our team managed and controlled our work activities. We tend to learn how to organize activities, determine their sequence, assign work, and monitor progress only with practice and experience. Such experiences are much more meaningful if we review and reconsider them after the fact. As with design tools, management tools are not equally useful in every setting—although some, such as the WBS, appear useful in almost every situation. In commercial settings, reviews of both budgets and work assignments provide critical grounding for future projects.

The last postproject audit step is a review of the project's outcome, in terms of the goals set and the processes used. While it is obviously useful to know whether or not we achieved our goals, it is important for us as team members to know whether this was a result of good planning, good execution, adequate resources, or simply good luck. In the long run, our team will have repeated successes only if we learn to plan well and to execute well.

Finally, we should not use a postproject audit to assign blame or to point fingers. Many project and institutional settings have formal mechanisms for peer review and supervisory evaluation of team members, and they provide valuable means for highlighting individual strengths, weaknesses, and contributions. But they also provide team members with important insights that they can use to improve their own work on design teams. That is why individual performance reviews are not central to, or even a desirable aspect of, a postproject audit. The audit is intended to show what the team and the organization did right to make the project successful, or what must be done differently in the future if the project was not successful.

## 16.7 NOTES

*Section 16.1*: The underlying conceptual model for the 3S approach is from Oberlander (2000).

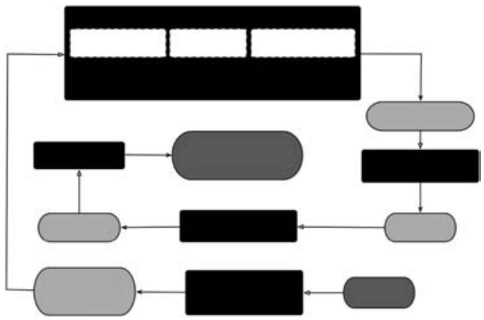
*Section 16.2*: The WBS for the Danbury project comes from Ahmad et al. (2007).

*Section 16.5*: The use of the standard form of the percent-complete matrix is given in Oberlander (2000).

# CHAPTER 17

# ETHICS IN DESIGN

*Design is not just a technical matter.*



DESIGN IS fundamentally a human endeavor, a *social activity*. The design process depends upon relationships among members within a design team, with clients and manufacturers, and with the purchasers and users of designed devices. In many cases, designs also affect people who were not part of the designer–client–user triangle we discussed in Chapter 1. To design means to *accept responsibility for creating designs*: designers are influenced by the society in which they work, and designed products influence society. That is why we must consider ethics and ethical behavior in our examination of how engineers design things.

## 17.1 ETHICS: UNDERSTANDING OBLIGATIONS

Words like ethics, morals, obligations, and duty are used in a variety of ways in everyday life, including seemingly contradictory or unclear ones. As we did with many of the engineering terms earlier in the book, we start with some definitions. First, the word *ethics*:

> **ethics 1:** the discipline dealing with what is good and bad and with moral duty and obligation **2 a:** a set of moral principles or values **b:** a theory or system of moral values **c:** the principles of conduct governing an individual or group

Since it is referenced so often in the definition of ethics, the word *moral*:

> **moral 1 a:** of or relating to principles of right or wrong in behavior **b:** expressing or teaching a conception of right behavior

Besides defining a discipline or field of study, these definitions define ethics as a set of guiding principles or a system that people can use to help them behave well. Most of us learn right and wrong from our parents, or perhaps as a set of beliefs from one of the religious traditions that emphasize faith in God (e.g., Christianity, Judaism, and Islam) or those that stress faith in a *right path* (e.g., Buddhism, Confucianism, and Taoism). However we learn about them, virtually all of us have a deep connection with notions such as honesty and integrity, and about the injunction to treat others as we would want to be treated ourselves.

If we already know these things, why do we need another, external set of rules? If we don't, and the law doesn't keep us in line, what is the use of a set of ethical principles? The answer is that the lessons we learn at home, in school, and in religious forums may not provide enough explicit guidance about many of the situations we face in life, especially in our professional lives. In addition, given the diversity and complexity of our society, it is helpful to have standards of professional behavior that are universally agreed upon, across all of our traditions and individual upbringings.

Our professional lives are also complicated because our responsibilities may involve obligations to many stakeholders, some of whom are obvious (e.g., clients, users, the immediately surrounding public) and some of whom are not (e.g., some government agencies, professional societies). We will elaborate these obligations later in this chapter, but we note for now that these obligations can conflict. For example, a client may want one thing, while a group of people affected by a design may want something different. Further, some of those people may not even know how they are being affected until *after* the design is complete and the design has been implemented. Just as we found in evaluating and ranking objectives during conceptual design, there is no simple formula or algorithm we can apply to decide among the various parties who may be influenced by our design. The priorities our clients, users, and the designer place on objectives are subjective in nature; so too is our personal assessment of the relative importance that we attach to our conflicting obligations. Professional codes of ethics provide a means of reconciling such competing obligations.

Consider the famous case in which group of engineers tried unsuccessfully to delay the launch of the space shuttle *Challenger* on January 28, 1986. While severe doubt was expressed by some engineers about the safety of the *Challenger*'s O-rings because of the cold weather before the flight, the upper management of Morton-Thiokol, the company that made the *Challenger*'s booster rockets, and NASA approved the launch. These managers, many of whom were themselves experienced engineers, determined that their concerns about Morton-Thiokol's image and the stature and visibility of NASA's shuttle program outweighed the judgments of other engineers who were closer to the booster design. The Morton-Thiokol engineers ultimately publicized the overruling of their recommendation not to launch by engaging in *whistleblowing*, where someone "blows the whistle" in order to stop a faulty decision made within a company, agency, or some other institution.

Whistleblowing is not new or unique. Another famous case is that of an industrial engineer, Ernest Fitzgerald, who spoke out on major cost overruns in the procurement of an Air Force cargo plane. The Air Force was so displeased with Fitzgerald's actions that it took actions to keep him from further work on the plane: It "lost" his Civil Service tenure and then reconstructed that part of the bureaucracy in which had Fitzgerald worked so as to eliminate his position! After an arduous and expensive legal battle, Fitzgerald earned a substantial settlement for wrongful termination and was reinstated in his position.

While such stories seem discouraging, they also show heroic behavior under trying circumstances. More to the point, these examples show how “doing what’s right” can be perceived quite differently within an organization. An engineer may be faced with a clash of obligations that lies at the crux of any discussion of engineering ethics. If that happens, to whom can the designer or engineer turn for help? While part of the answer lies in the engineer’s personal understanding of ethics, another part of the answer lies in the support of professional colleagues and peers. One of the primary sources of such insight and guidance is the professional engineering societies and their codes of ethics.

## 17.2 CODES OF ETHICS: WHAT ARE OUR PROFESSIONAL OBLIGATIONS?

Imagine you are a mining engineer who has been engaged by the owner of a mine to design a new shaft extension. As part of that design task you survey the mine and find that part of it runs under someone else’s property. Are you obligated to simply complete the survey and the design for the mine owner who is paying you, and then go on to your next professional engagement?

Suppose you suspect that the mine owner hasn’t notified the landowner that his mineral rights are being excavated out from underneath him. Should you do something about that? If so, what? Further, what compels you to do something? Is it personal morality? Is there a law? How are you responsible, and to whom?

The chain of questions just started can easily be lengthened, and the situation made more complicated. For example, what if the mine was the only mine in town and its owner controlled your livelihood, and those of many town residents? If you find out that the mine runs under an elementary school, does that change things?

This story highlights some of the actors and obligations that could arise in an engineering project. In fact, scenarios such as this occurred late in the 19th century and early in the 20th; such situations provided part of the impetus to the formation of professional societies and the development of codes of ethics as a form of protection for their members.

Over time, the professional societies also undertook other kinds of activities, including promulgating design standards, and providing forums for reporting research and innovations in practice. The professional engineering societies continue to play a leading role in setting ethical standards for designers and engineers. These ethical standards clearly speak to the various and often conflicting obligations that an engineer must meet. The societies also provide mechanisms for helping engineers resolve conflicting obligations, and, when asked, they provide the means for investigating and evaluating ethical behavior.

Most professional engineering societies have published codes of ethics. We show the codes of ethics of the American Society of Civil Engineers (ASCE) in Figure 17.1 and of the Institute of Electronics and Electrical Engineers (IEEE) in Figure 17.2. While both codes emphasize integrity and honesty, they do appear to value certain kinds of behavior differently. For example, the ASCE code enjoins its members from competing unfairly with others, a subject not mentioned by the IEEE. Similarly, the IEEE specifically calls for its members to “fairly treat all persons regardless of such factors as race, religion, gender . . . ” There are

## ASCE CODE OF ETHICS

### Fundamental Principle

Engineers uphold and advance the integrity, honor, and dignity of the engineering profession by:

1. using their knowledge and skill for the enhancement of human welfare and the environment;
2. being honest and impartial and serving with fidelity the public, their employers and clients;
3. striving to increase the competence and prestige of the engineering profession; and
4. supporting the professional and technical societies of their disciplines.

### Fundamental Canons

1. Engineers shall hold paramount the safety, health and welfare of the public and shall strive to comply with the principles of sustainable development [1] in the performance of their professional duties.
2. Engineers shall perform services only in areas of their competence.
3. Engineers shall issue public statements only in an objective and truthful manner.
4. Engineers shall act in professional matters for each employer or client as faithful agents or trustees, and shall avoid conflicts of interest.
5. Engineers shall build their professional reputation on the merit of their services and shall not compete unfairly with others.
6. Engineers shall act in such a manner as to uphold and enhance the honor, integrity, and dignity of the engineering profession and shall act with zero-tolerance for bribery, fraud, and corruption.
7. Engineers shall continue their professional development throughout their careers, and shall provide opportunities for the professional development of those engineers under their supervision.

[1] In November 1996, the ASCE Board of Direction adopted the following definition of Sustainable Development: "Sustainable Development is the challenge of meeting human needs for natural resources, industrial products, energy, food, transportation, shelter, and effective waste management while conserving and protecting environmental quality and the natural resource base essential for future development."

**FIGURE 17.1** The code of ethics of the ASCE, as modified July 2006. It is similar, although not identical, to the code adopted by the IEEE that is displayed in Figure 17.2.

other differences as well in the styles of language. The ASCE presents a set of injunctions about what engineers "shall" do, while the IEEE code is phrased as a set of commitments to undertake certain behaviors.

Notwithstanding these differences, both codes of ethics set out guidelines or standards of how to behave with respect to: clients (e.g., ASCE's "as faithful agents or trustees"); the profession (e.g., IEEE's "assist colleagues and coworkers in their professional development"); the law (e.g., IEEE's "reject bribery in all its forms"); and the public (e.g., ASCE's "shall issue public statements only in an objective and truthful manner"). Perhaps most noteworthy, both place a primary concern on protection of the health, safety, and welfare of the public. We will return to this paramount principle in Section 17.5.

The codes of ethics, along with the interpretations and guidance offered by the societies, lay out "rules of the road" for dealing with conflicting obligations, including the task of assessing whether these conflicts are "only" of perception or of a "real" and potentially damaging nature.

There are some points to make regarding the professional societies and their codes. First, the differences in the codes reflect different styles of engineering practice in the various disciplines much more than differences in their views of the importance of ethics.

**IEEE CODE OF ETHICS**

We, the members of the IEEE, in recognition of the importance of our technologies in affecting the quality of life throughout the world, and in accepting a personal obligation to our profession, its members and the communities we serve, do hereby commit ourselves to the highest ethical and professional conduct and agree:

1. to accept responsibility in making decisions consistent with the safety, health and welfare of the public, and to disclose promptly factors that might endanger the public or the environment;
2. to avoid real or perceived conflicts of interest whenever possible, and to disclose them to affected parties when they do exist;
3. to be honest and realistic in stating claims or estimates based on available data;
4. to reject bribery in all its forms;
5. to improve the understanding of technology, its appropriate application, and potential consequences;
6. to maintain and improve our technical competence and to undertake technological tasks for others only if qualified by training or experience, or after full disclosure of pertinent limitations;
7. to seek, accept, and offer honest criticism of technical work, to acknowledge and correct errors, and to credit properly the contributions of others;
8. to treat fairly all persons regardless of such factors as race, religion, gender, disability, age, or national origin;
9. to avoid injuring others, their property, reputation, or employment by false or malicious action; and
10. to assist colleagues and co-workers in their professional development and to support them in following this code of ethics.

**FIGURE 17.2** The code of ethics of the IEEE, dated February 2006. How does the IEEE code of ethics differ from that adopted by the ASCE that is displayed in Figure 17.1?

For example, most civil engineers who are not employed by a government agency work in small companies that are people-intensive, rather than capital-intensive. These firms obtain much of their work through public, competitive bidding. Electrical engineers, on the other hand, typically work in the private sector for corporations that sell products more than services, one result of which is that they have significant manufacturing operations and are capital-intensive. Such different practices produce different organizational cultures and, hence, different ways of expressing ethical standards.

A second point is that the professional societies, notwithstanding their promulgation of codes of ethics, have not always been seen as active and visible protectors of whistleblowers and other professionals who raise concerns about specific engineering or design instances. This situation is improving, steadily if slowly, but many engineers still find it difficult to look to their societies, and especially their local branch sections, for first-line assistance and support in times of need. Of course, as we all make ethical behavior a more visible priority, the need for such support may lessen, and its ready availability will surely increase.

Finally, we should note that the codes of ethics we have described are those found in the United States and Canada, which are not necessarily the same as those in other parts of the world. In countries with a strongly Islamic culture and government, for example, the codes of ethics often reflect an alignment between religious values and professional practice that seems alien to our tradition of separating church and state. Similarly, the code of ethics of the *Verein Deutscher Ingenieure* (VDI), or Association of German Engineers, is based on the historical need of German engineers to reflect upon and respond to the willingness of many of their colleagues to work in support of the Nazis during the 1930s and 1940s. It is important for us, as professional engineers, to understand and respond to the culture in which we work, while remaining true to our own values.

## 17.3 OBLIGATIONS MAY START WITH THE CLIENT . . .

Let us consider in greater depth our obligations to a client or to an employer. As designers or engineers, we owe our client or employer a professional effort at solving a design problem, by which we mean being technically competent, conscientious, and thorough, and that we should undertake technical tasks only if we are properly "qualified by training or experience." We must avoid any conflicts of interest, and disclose any that may exist. And we must serve our employer by being "honest and impartial" and by "serving with fidelity . . . " Most of these obligations are clearly delineated in codes of ethics (e.g., compare the quotes with Figures 17.1 and 17.2), but there's at least one curious obligation on this list: What does it mean to "serve with fidelity"?

A thesaurus would tell us that fidelity has several synonyms, including constancy, fealty, allegiance, and loyalty. Thus, one implication we can draw from the ASCE code of ethics is that we should be loyal to our employer or our client. This suggests that one of our obligations is to look out for the best interests of our client or employer, and to maintain a clear picture of those interests as we do our design work. But loyalty is not a simple, one-dimensional attribute. In fact, clients and companies earn the loyalty of their consultants and staffs in at least two ways. One, called *agency-loyalty*, derives from the nature of any contracts between the designer and the client (e.g., "work for hire") or between the designer and her employer (e.g., a "hired worker"). As it is dictated by contract, agency-loyalty is clearly obligatory for the designer. A second kind of loyalty, *identification-loyalty*, is more likely to be perceived as optional. It stems from aligning oneself with the client or company because the engineer admires its goals or sees its behavior as mirroring his or her own values. To the extent that identification-loyalty is optional, it will be earned by clients and companies only if they reciprocate by demonstrating loyalty to their own staff designers.

Agency-loyalty provides one reason to maintain a "design notebook" to document design work. As we have noted before, keeping such a record is good design practice because it is useful for recapitulating our thinking as we move through different stages of the design process and for real-time tracking. A dated design notebook also provides legal documentation of how and when new patentable ideas were developed. Such documentation is essential to an employer or client if a patent application is in any way challenged. Further, as is typically specified in contracts and employment agreements, the intellectual work done in the creation of a design is usually the intellectual property of the client or the employer. A client or an employer may share the rights to that intellectual property with its creators, but the fundamental decisions about the ownership of the property generally belong to the client or owner. It is important for a designer to keep that in mind, and also to document any separate, private work that she is doing, just to avoid any confusion about who owns any particular piece of design work.

Since identification-loyalty is optional, it provides fertile ground for clashes of obligations because other loyalties have the space to arise. As we discuss further in Section 17.5, modern codes of ethics normally articulate some form of obligation to the health and welfare of the public. For example, the ASCE code of ethics (viz., Figure 17.1) suggests that civil engineers work toward enhancing both human welfare and the environment, and that they "shall hold paramount the safety, health, and welfare of the public . . . " Similarly, the IEEE code (viz., Figure 17.2) suggests that its members commit to "making engineering decisions consistent with the safety, health, and welfare of the public . . . " These are clear calls to engineers to identify and honor loyalties beyond their employer. It was just such divided loyalties that emerged in the cases of whistleblowing discussed above.

In the case of the explosion of the *Challenger*, those who argued against its launching felt that lives would be endangered. They placed a higher value on the lives being risked than they did on the loyalties being demanded by Morton-Thiokol (i.e., to secure its place as a government contractor) and by NASA (i.e., to its ability to successfully argue for the shuttle program before Congress and the public). Similar conflicting loyalties between employer and the public emerged for engineers as toxic waste sites were cleaned up under the Superfund program of the Environmental Protection Agency (EPA). There are other cases in which engineers were apparently willing to rank their loyalties to their companies first, to the point where falsified test data were reported to the government (by engineers and managers at the Ford Motor Company), or parts known to be faulty were delivered to the Air Force (by engineers and managers at the B. F. Goodrich Company).

An apparent disloyalty to a company or an organization may sometimes be, in a longer term, an act of greater, successfully merged loyalties. When the Ford Pinto was initially being designed, for example, some of its engineers wanted to perform crash tests that were not required by the relevant U.S. Department of Transportation regulations. The managers overseeing the car's development felt that such tests could not benefit the program and, in fact, might only prove to be a burden. Why run a test that is not required, only to risk failing that test? The designers who proposed the tests were seen as disloyal to Ford and to the Pinto program. In fact, the placement of the drive train and the gas tank resulted in fiery crashes, lives lost, and major public relations and financial headaches for Ford. Clearly, Ford would have been better off in the long run to have conducted the tests, so the engineers who proposed them could be said to have been looking out for the company's long-term interests.

If there is one point that emerges from the discussion thus far, it is that ethical issues do not arise from a *single* obligation. Indeed, were issues so easily categorized, choices would vanish and there would be no ethical conflicts. The very existence of professional codes of ethics, however, testifies to the reality of conflicting obligations and provides guidance to mediating those conflicts.

## 17.4 . . . BUT WHAT ABOUT THE PUBLIC AND THE PROFESSION?

When people behave responsibly, things can turn out well even in bad situations, which we will demonstrate with a true story. In fact, to start with the ending, the protagonist-hero of this story said, "In return for getting a [professional engineering] license and being regarded with respect, you're supposed to be self-sacrificing and look beyond the interests of yourself and your client to society as a whole. And the most wonderful part of my story is that when I did it nothing bad happened."

Our hero is the late William J. LeMessurier (pronounced "LeMeasure") of Cambridge, Massachusetts, one of the most highly regarded structural engineers and designers in the world. He served as the structural consultant to a noted architect, Hugh Stubbins, Jr., for the design of a new headquarters building for Citicorp in New York City. Completed in 1978, the 59-story Citicorp Center is still one of the most dramatic and interesting skyscrapers in a city filled with some of the world's great buildings (see Figure 17.3). In many ways, LeMessurier's conceptual design for Citicorp resembles other striking

Photo by Clive L. Dym

**FIGURE 17.3** One view of the 59-story Citicorp Center, designed by architect Hugh Stubbins, Jr., with William J. LeMessurier serving as structural consultant. One of this building's notable features is that it rests on four massive columns that are placed at the midpoints of the building's sides, rather than at the corners. This enabled the architects to include under the Citicorp's sheltering canopy a new building for St. Peter's Church.

skyscrapers in that it used the *tube* concept in which a building is designed as a tall, hollow tube that has a comparatively rigid or stiff tube wall. (In structural engineering terminology, the tube's main lateral stability elements are located at the outer perimeter and tied together at the corners.) Fazlur Kahn's John Hancock Center in Chicago is a similar design (see Figure 17.4). The outer "tube" or "main lateral stability elements" are the multistory diagonal elements that are joined to large columns at the corners. Kahn's design benefited from a deliberate architectural decision to expose the tube's details, perhaps to illustrate the famous dictum that *form follows function* (which originated with Louis Sullivan, the noted Chicago architect who was Frank Lloyd Wright's mentor).

LeMessurier's Citicorp design was innovative in several ways. One, not visible from the outside, was the inclusion of a large mass, floating on a sheet of oil, within the triangular roof structure. It was added as a damper to reduce or damp out any oscillations the building might undergo due to wind forces. Another innovation was LeMessurier's adaptation of the tube concept to an unusual situation. The land on which the Citicorp Center was built had belonged to St. Peter's Church, with the church occupying an old Gothic building on the lot's west side. When St. Peter's sold the building lot to Citicorp, it also negotiated that a new church be erected "under" the Citicorp skyscraper. In order to manage this, LeMessurier moved the "corners" of the building to the midpoints of each side (see Figure 17.5). This enabled the creation of a large space for the new church because the office tower itself was then cantilevered out over the church, at a height of some nine stories. Looking at the sidewalls of the tube—and here we have to peel off the building's skin because architect Stubbins did not want the structures exposed as they were in the Hancock tower—we see that the wall's rigidity comes from large triangles, made up of diagonal and horizontal elements, that all connect at the midpoints of the sides. Thus, LeMessurier's triangles serve the same purpose as Kahn's large X-frames.

Photo by Clive L. Dym

**FIGURE 17.4** The 102-story John Hancock Center, designed by the architectural firm of Skidmore, Owings and Merrill, with Fazlur Kahn serving as structural engineer. Note how the exposed diagonal and column elements make up the tube that is the building's underlying conceptual design.

The ethics problem arose soon after the building was completed and occupied. LeMessurier received a call from an engineering student who was told by a professor that the building's columns had been put in the wrong place. LeMessurier was very proud of his idea of placing the columns at the midpoints. He explained to the student how the 48 diagonal braces he had superposed on the midspan columns added great stiffness to the building's tube framework, particularly with respect to wind forces. The student's questions sufficiently intrigued LeMessurier that he later reviewed his original design and calculations to see just how strong the wind bracing system would be. He found himself looking at a case that was not covered under then-current practice and building codes. Practice at the time called for wind force effects to be calculated when the wind flow hit a side of a building dead on, that is, normal to the building faces. However, the calculation of the effect of a *quartering wind*, under which the wind hits a building on a 45° diagonal and the resulting wind pressure is then distributed over the two immediately adjacent faces (see Figure 17.6), had not been called for previously. A quartering wind on the Citicorp Center leaves some diagonals unstressed and others doubly loaded, with calculated strain increases of 40%. Normally, even this increase in strain (and stress) would not have been a problem because of the basic assumptions under which the entire system was designed.

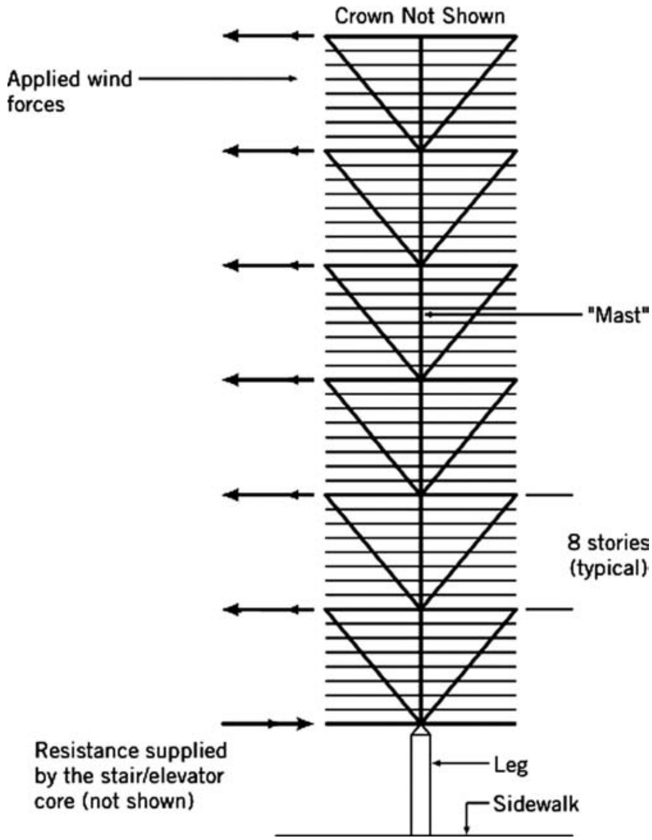



**FIGURE 17.5** A sketch of LeMessurier's Citicorp design. Here the tube is made up of (unexposed) diagonal elements, organized as rigid triangles, and connected to the four columns at the midpoints of the sides of the building. Adapted from *Civil Engineering*.

However, LeMessurier learned a few weeks later that the actual connections in the finished diagonal bracing system were not the high-strength welds he had stipulated. Rather, the connections were bolted because Bethlehem Steel, the steel fabricator, had determined and suggested to LeMessurier's New York office that bolts would be more than strong enough and, at the same time, significantly cheaper. The choice of bolts was sound and quite correct professionally. However, to LeMessurier the bolts meant that the margin of safety against forces due to a quartering wind—which, again, structural engineers were not then required to consider—was not as large as he would have liked. (It is interesting to note that while New York City's building code did not then require that quartering winds be considered in building design, Boston's code did—and had since the 1950s!)

Spurred by his new calculations and the news about the bolts, and by hearing of some other detailed design assumptions made by engineers in his New York office, LeMessurier retreated to the privacy of his summer home on an island in Maine to carefully review all of the calculations and changes and their implications. After doing a member-by-member calculation of the forces and reviewing the weather statistics for New York City,

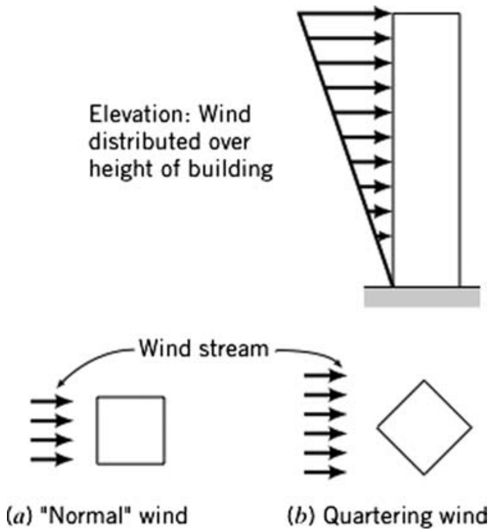



**FIGURE 17.6** A sketch of how wind forces act on buildings. (a) This is the standard case of wind streams being normal or perpendicular to the building faces. It was the only one called for in design codes at the time the Citicorp center was designed. (b) This is the case of quartering winds, in which case the wind stream come along a 45° diagonal and thus simultaneously applies pressure along two faces at once.

LeMessurier determined that statistically, once every 16 years the Citicorp Center would be subjected to winds that could produce a catastrophic failure. Thus, in terminology used by meteorologists to describe both winds and floods, the Citicorp Center would fail in a 16-year storm—when it had supposedly been designed to withstand a 50-year storm. So, what was LeMessurier to do?

In fact, LeMessurier considered several options, reportedly including driving into a freeway bridge abutment at high speed. He also considered remaining silent, as he tried to reassure himself that his innovative rooftop mass damper actually reduced the probabilities of such failure to the 50-year level. On the other hand, if the power went out, the mass damper wouldn't be there to help. So, what did LeMessurier actually do?

He first tried to contact the architect, Hugh Stubbins, who was away on a trip. He then called Stubbins' lawyer, after which he talked with his own insurance carrier and then with the principal officers of Citicorp, one of whom had studied engineering before choosing to become a banker. While some early consideration was given to evacuating the building, especially since hurricane season was just over the horizon, it was decided instead that all of the connections at risk should be redesigned and retroactively fixed. Steel plate "Band-Aids," 2 in. thick, would be welded onto each of 200 bolted connections. However, there were some very interesting implementation problems, only a few of which we mention here. The building's occupants had to be informed without being alarmed, because the repair work would go on every night for 2 months or more. The public had to be informed why the bank's new flagship headquarters suddenly needed immediate modifications. (In fact, the entire process was open and attentive to public concerns.) Skilled structural welders, who were in short supply, had to be found, as did an adequate supply of the right grade of steel plate. Discrete, even secret, evacuation plans had to be put in place, just in case an unexpected storm came up while the repairs were being made. And New York City's Building Commissioner and the Department of Buildings and its inspectors had to be brought into the loop because they were central to resolving the problem. They had to be informed about the problem and its proposed solution, and they had to agree to inspect that solution. All in all, a dazzling array of concerns, institutions, and, of course, personalities.

In the end, the steel Band-Aids were applied and the entire business was completed professionally, with no finger pointing and no public assignments of blame. LeMessurier, who had thought his career might precipitously end, came away with still greater stature, occasioned by his willingness to face up to the problem candidly and propose a realistic, carefully crafted solution. In the words of one of the engineers involved in implementing LeMessurier's solution, "It wasn't a case of 'We caught you, you skunk.' It started with a guy who stood up and said, 'I got a problem, I made the problem, let's fix the problem.' If you're gonna kill a guy like LeMessurier, why should anybody ever talk?"

This is a case where everyone involved behaved well. In fact, it is to everyone's credit that all of the participants acted with a very high standard of professionalism and understanding. It is, therefore, a case that we can study with pleasure, particularly as engineers. It is also a case that could have gone other ways, so we will close our discussion by posing a few questions that you, the reader, might face were you in LeMessurier's position:

- Would you have "blown the whistle," or not?
- What would you have done if you determined that the revised probability of failure was higher (i.e., worse) than for the original design, but still within the range permitted by code?
- What would you have done if your insurance carrier had said to "keep quiet"?
- What would you have done if the building owner, or the city, had said to "keep quiet"?
- Who should pay for the repair?

## 17.5 ON ENGINEERING PRACTICE AND THE WELFARE OF THE PUBLIC

It is easy to imagine a scenario where we are asked to design a product that we think needn't be made, or perhaps even shouldn't be made. Earlier in this book, for example, we referred to the design of a cigarette lighter, which we also thought of as an igniter of leafy matter. While this example seems trivial, even nonsensical, it points to another facet of divided loyalties. It suggests that designing cigarette lighters might somehow be morally troubling. In the United States today, there are many people who would consider designing cigarette lighters and cigarette-making machinery as being at least "politically incorrect," and perhaps even morally wrong. On the other hand, isn't it up to individuals to choose to smoke or not? If a product is legal, should we allow ourselves to design it without feeling uncomfortable? How should the effect on users and the public influence our choice of project or client?

A much more serious instance that profoundly extends this line of thought emerges when we consider the design of large-scale ovens and associated specialized buildings made in Nazi Germany in the 1930s and 1940s. Or we might consider also the design of nuclear weapons in the United States or the former Soviet Union in the 1950s–1990s, and today in a growing number of developing countries. Were these designers merely being loyal to their clients, their governments, and to their societies? If so, how does this reconcile with a commitment to "human welfare and the environment"?

Recall that the codes of ethics we discussed in Section 17.2 place the health, safety, and welfare of the public in the first or paramount position. Historically, most engineers and the professional societies have focused almost entirely on the health and safety aspects of these phrases. In a manner similar to the medical profession's admonition, "First, do no harm," engineers are committed to ensuring that the things we design are not willfully dangerous, and that the process of design is rigorous, thorough and honest about potential risks to the public. Unfortunately, the "welfare of the public" phrase has not been so deeply explored or considered. Some philosophers of technology have challenged engineers to consider these issues more carefully as well.

For most of us, concern with human welfare begins with meeting fundamental human needs, such as ensuring adequate food, water, and shelter. While starting from basic needs, and perhaps taking a cue from contemporary economics, we often extend these concerns so that "more" and "better" become synonyms. They certainly are the same for those around the world who live in abject poverty. It is, however, an open question whether "more" and "better" are the same in the developed world. Clearly, when we speak of the "welfare of the public," we are no longer in the land of the purely technical: The welfare of the public is implicitly about what constitutes "the good life." The depletion of key resources, the degradation of our environment, and the changes in our global climate should give us pause about what we mean by "the good life."

The phrase "the welfare of the public" suggests that we, as engineering designers, must be aware that our work can move into the social and even political realm. For example, the ASCE has issued Guidelines to Practice that state that in order to adhere to the canons of its code of ethics (Figure 17.1) engineers should "recognize that the lives, safety, health, and welfare of the general public are dependent upon engineering judgments, decisions and practices incorporated into structures, machines, products, processes, and devices." This guideline clearly frames the *judgments*, *decisions*, and *practices* as more than just technical or scientific aspects of engineering. It explicitly links them to the well-being of the public. Thus, the practice of engineering has much in common with other professions such as law and medicine: Engineers bear a special responsibility to practice their profession with an awareness of the larger environment.

Engineers must exercise a sort of social judgment when doing design because of the results that arise *after* design, once a device or system is released for public use. Technology is often not neutral—it has effects and consequences, only some of which can be anticipated and controlled. Technology is implicated in particular possibilities for social organization and social relationships, in the establishment and enforcement of permissions and prohibitions, and in the distribution of economic, social, and political power. Andrew Feenberg articulated this insight most clearly in his book *Questioning Technology*:

> *Technology is power in modern societies, a greater power in many domains than the political system itself. The masters of technical systems, corporate and military leaders, physicians and engineers, have far more control over patterns of urban growth, the design of dwellings and transportation systems, the selection of innovations, our experience as employees, patients and consumers, than all the electoral institutions of our society put together. But, if*

*this is true, technology should be considered as a new kind of legislation, not so very different from other public decisions. The technical codes that shape our lives reflect particular social interests to which we have delegated the power to decide where and how we live, what kinds of food we eat, how we communicate, are entertained, healed and so on.*

One of the more challenging questions facing engineers who take the welfare of the public seriously is "Who, or what, is the public?" When an engineer begins a project to design a water treatment facility, she needs to be attentive to the complexity of the public whose well-being she has committed to serve by participating in her profession. Engineers can go a long way toward satisfying this need simply by taking it seriously. This means that attending to the public interest requires us to practice an ongoing, good-faith effort to listen to affected parties and those who speak for them.

As engineers, we should realize that what we design *creates* publics. The engineer who proposes a highway to a group of homeowners may find to his surprise that he has created a public (and energized its opposition to the project) without ever intending to do so. Similarly, no one anticipated the scope or reach of the Internet, yet its designers surely created many "publics," including some who used web technologies to help overthrow dictators.

## 17.6 ETHICS: ALWAYS A PART OF ENGINEERING PRACTICE

Ultimately, ethics is intensely personal. Returning to our question about designing cigarette lighters, the matter necessarily resolves to, should *I* be working on this design project? While the professional societies produce and insist upon standards of professional conduct, it is individual practitioners who practice engineering. There is no way to predict when a serious conflict of obligations and loyalties will arise in our individual lives. Nor can we know the specific personal and professional circumstances within which such conflicts will be embedded. Nor, unfortunately, is there a single answer to many of the questions posed. If faced with a daunting conflict, we can only hope that we are prepared by our upbringing, our maturity, and our ability to think and reflect about the issues that we have briefly raised here.

## 17.7 NOTES

*Section 17.1*: Martin and Schinzinger (1996) and Glazer and Glazer (1989) are very interesting, useful, and readable books on, respectively, engineering ethics and whistleblowing. Ethics emerges as a major theme in Harr (1995) tale of a civil lawsuit spawned by inadequate toxic waste cleanup.

*Section 17.2*: An interesting account of the historical development of the professional societies and codes of ethics if given in Davis (1992). The issues relating to international codes of ethics are discussed in Little et al. (2008).

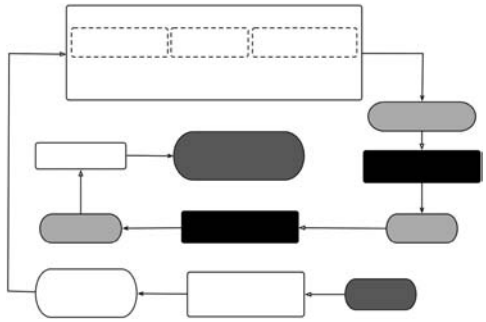
*Section 17.3*: The definitions of agency- and identification-loyalty are derived from Martin and Schinzinger (1996).

*Section 17.4*: The Citicorp Center case is adapted from Morgenstern (1995) and Goldstein and Rubin (1996). We were greatly helped by William LeMessurier's review of the material.

*Section 17.5*: This section raises issues that require careful and thoughtful reading. For the sake of brevity, we cite only Arendt (1963), Harr (1995), Feenberg (1990), Little et al. (2008), and Winner (1990) among the many sources about the deep and complex issues raised in this section.

APPENDIX A

# PRACTICAL ASPECTS OF PROTOTYPING



THERE ARE several important practical matters to think about when prototyping. First and foremost, before working in a machine or woodworking shop, we need to prepare by learning basic shop safety lessons. We thus start this appendix with a brief overall summary of shop safety. Then we discuss how we select materials, build things, and choose fasteners, as we build models and prototypes.

## A.1 WORKING SAFELY IN A SHOP

Safety in the workshop is critically important. Power tools can easily cause dismemberment and death. A moment's inattention can lead to a permanent change in lifestyle and career. *Take these safety warnings seriously*:

- Do not use equipment or machinery for which you have not been trained.
- Use protective gear and dress properly.
- Do not use power or machine tools when you are tired or intoxicated.
- Always have a buddy with you in the shop.

Almost all machine tools come with brochures that detail their safe use. *Read that documentation*. The time you spend may save you a finger or an eye. Most facilities have training programs or videos available, so be sure to ask for them and use them. In fact, you are often not allowed to use machinery or shops until you have passed the safety training. *Do not try to bypass or cheat on the training*. The main rule is: *Keep your body* *parts away from sharp moving objects*. There is almost always a safe way to do something. Learn what that safe way is.

Individual shops will have their own safety requirements for dress and protective gear. Consider the following to be a minimal set of requirements:

- *Always wear eye protection, safety goggles, or safety glasses, while using tools or when you are near someone who is*. Drills, lathes, mills, and saws produce shavings that often become airborne. Hammers do shatter, and objects being struck by hammers often break or go flying off. Wrenches and screwdrivers tend to be less hazardous to eyes, but freak accidents do occur. *Keep your eyes safe*.
- *Keep your hair short or pulled back and out of the way.* Spinning drills, mills, and lathes seem to have a magnetic attraction for long hair, and it can easily get trapped by a rapidly spinning drill chuck or a speeding saw band. Hair pulled back into a ponytail still presents a hazard if the hair can fall forward into a machine. *Keep your hair intact*.
- *Always wear full-coverage shoes with a sturdy sole*. Sandals or flip-flops will not protect you from dropped tools or hot metal shavings. Thin soles will not protect you from sharp objects found on the floor. *Protect your feet*.
- *Wear long, non-baggy pants*. As with shoes, long pants will protect your legs against hot flying metal shavings and other hazards. Baggy pants can get caught in moving or rotating machinery. Pants don't have to be tight, but the closer they fit, the less chance you have of getting them caught on something.
- *Wear short-sleeved shirts or blouses*. Many serious injuries and fatalities are caused when loose clothing gets caught in moving machinery. Short sleeves represent a trade-off between protecting an arm and getting a sleeve caught in a drill press or a mill. Rolled-up long sleeves don't count because they can unroll and get tangled in the machinery. Any loose fabric that is within a foot of your hands is likely to get caught in a machine tool that you are using.
- *Do not wear jewelry around machine tools*. Take the jewelry off and store it somewhere safe. Necklaces, watches, and bracelets are the most dangerous items, but rings, earrings, and other piercings can get caught and do damage.

Check that there is proper protection against fumes or particulates. Make sure ventilation is adequate, and that you meet requirements regarding particulates. Fatigue and intoxication cause many industrial accidents. It is far better to miss a deadline (or risk a lower grade!) and keep all of your fingers than it is to miss the deadline (or get the lower grade) anyway because you had to make an unplanned trip to the emergency room or spend time in the hospital!

## A.2 SELECTING MATERIALS

There are many choices for a material to use to building a prototype or model. Here we review a few of the most common materials and their basic properties to provide a basis for choosing appropriate materials.

*Paper and cardboard* are suitable for inexpensive models that do not have to support large loads. They are also good stand-ins when a properly equipped workshop is not available. Paper and cardboard are usually measured and marked with rulers, pencils, compasses, and templates or stencils. They are usually cut with scissors, paper cutters, or knives such as box cutters or hobby knives. They are usually folded or rolled into their final shapes, and fastened with glue, paste, tape, staples, or roundhead fasteners.

*Wood* is usually purchased as lumber. Lumber comes from a number of different trees, each with its own properties, and may come green or dried. Green wood is lumber that has not been allowed to dry after being cut to shape and will change in dimension quite drastically with time. Dried lumber, especially kiln-dried, has had much of the moisture removed from it and will remain much more dimensionally stable. Lumber is also classified as softwood (such as Douglas fir, pine, or redwood), or hardwood (such as oak, cherry, or walnut). The nominal sizes of softwood lumber range from 1 in. $\times$ 2 in. to 8 in. $\times$ 8 in. The most common nominal size is 2 in. $\times$ 4 in. Standard lengths range from 4 to 16 ft. The true dimensions are usually less than the nominal and vary with water content. A dried softwood 2 in. $\times$ 4 in. is actually closer to $1\frac{1}{2}$ in. $\times$ $3\frac{1}{2}$ in. The nominal length is usually close to the true length. Hardwoods are also available in fractional sizes from 3/8 in. on up. The nominal sizes of hardwoods are usually a little closer to the true dimensions. Softwoods are much less expensive than hardwoods, while hardwoods tend to be stronger and more wear-resistant.

Wood is an anisotropic material, that is, the properties along the grain differ greatly from those across the grain. The tensile strength is much greater along the grain than across. Thus we have to think carefully about stress directions when using wood. For example, beams carry their loads by developing stresses in the direction normal to their given applied loads.

Wood can be cut with assorted handsaws and power saws, including band saws and scroll saws. It can be shaped with a wood lathe, a router, or a power sander. It can be drilled with either a hand drill or a drill press.

Wood dimensions and flatness will vary with humidity and exposure to water or other absorbable fluids. It is very difficult to maintain tight dimensional tolerances in wood. Expansion spaces have to be designed into closed wood structures (such as boxes or drawers) so that undue stress is not placed in the wood as it expands and contracts in response to the weather.

*Plywood* is a composite material that is made from thin layers of wood glued together. There are a large number of different grades and thicknesses. The standard size sheet is 4 ft $\times$ 8 ft. Plywood is much more dimensionally stable than wood, and much more uniform in properties. However, it is also anisotropic, being much stronger in the plane of the layers than it is in a direction normal to the layers. It can be shaped with the same tools used for wood, but will cause faster wear on the cutting tools.

*Polymers* or *plastics* such as PVC, ABS, polystyrene, and acrylic are available in a number of preformed shapes, including sheets, bars, strips, film, rods, disks, tubes and pipes, and U-channels. Most of the polymers can be shaped with the same tools as wood. They can also be cut or shaped on the same machine tools used for metals if the tool's speed is suitably adjusted. Polymers can hold tight tolerances quite well and, depending on the polymer, can be machined into quite intricate shapes. There are usually solvent-based adhesives for joining one piece of a polymer to another piece of the same polymer. If properly done, the joint has the same strength as the bulk material. Polymers are not as strong or stiff as aluminum or steel, but they can be quite strong in some applications.

*Aluminum* is available in a large number of grades and shapes. Bulk aluminum is available as sheets, bars, strips, film, rods, disks, tubes, and U-channels, among other shapes.

Aluminum is quite strong and light. It is not as strong as steel, but it is easier to shape and machine. It holds dimensional tolerances very well. It can be cut with a band saw or hacksaw, machined with metal lathes or mills, and holes can be drilled with a mill, drill press, or hand drill. Aluminum has very high thermal and electrical conductivities. If not brought into contact with iron or steel, aluminum is quite resistant to corrosion at room temperature. Aluminum melts at a fairly low temperature and is not suitable for parts that will be exposed to high temperatures. Certain aluminum alloys can be welded, but it requires specialized tools. Aluminum is best joined with fasteners such as machine screws or rivets.

*Mild steel* is denser and stronger than aluminum. Bulk steel is also available as sheets, bars, strips, film, rods, disks, tubes, U-channels, and other shapes. It can also be cut with a band saw or hacksaw, machined with metal lathes or mills, and holes can be drilled with a mill, drill press, or hand drill. However, the cutting tools will experience increased wear with steel and the cutting speeds will be lower. Steel does require protection against corrosion. Perhaps the major use of steel in hand-built prototypes is in sheet form, and sheet steel is easily spot welded (see Section A.5).

The following are the two final details on materials: First, the choice of materials will be governed by cost, model performance requirements (e.g., are we building a true prototype or a model?), and access to cutting and shaping tools. Second, we will also need fasteners made of appropriate materials in order to connect the parts of our model or prototype together. Fasteners include things such as nails, wood screws, machine screws, sheet metal screws, bolts, nuts, washers, and pins. We will now detail the process of selecting fasteners and the techniques for properly using and installing a fastener.

## A.3 BUILDING TECHNIQUES

This section describes some of the basic techniques for shaping and joining materials. There are many sites on the Web and in reference libraries that provide further information.

*Straight edges* in wood and polymers are usually best cut with a table saw or a band saw that has a guide rail. The rail is set at the required distance from the blade, the piece is held firmly against the guide rail, and the piece is then pushed through the saw. Be sure to use a push stick or rod if necessary to keep your fingers a safe distance away from the blade. Straight edges in metal are usually rough cut with a band saw or cutting wheel, and then finished or faced with a mill.

*Curved edges* in wood or plastic are usually cut with a band saw or scroll saw. The desired profile is drawn in pencil on the wood and then the pencil line is used to guide the saw. Care must be taken not to bind the saw blade (put lateral forces on it) when cutting curves. The curve can be hand or power sanded to smooth off the saw cut. Curved edges in metal may be cut with a proper band saw. The blade should be appropriate for the metal being cut and a great deal of patience will be required, because it will take much longer to cut metal than it would a similar thickness of wood. The curve can be filed or hand- or power-sanded to a final shape.

*Cylindrically symmetric profiles* can be formed in wood on a wood lathe, and in metal and polymers on a metal lathe. Since the operation of lathes and similar tools is beyond our scope, be sure to obtain the proper training if you feel that use of such tools would be appropriate to your task.

*Holes* are used to allow passage for a fastener or for something like an axle, a cable, or a tube. In mass-produced items, the locations of such *joiner holes* are specified with geometric tolerances that determine whether parts are easily interchangeable. In models and one-offs, it is usually quicker and easier to clamp two parts together and drill through both at once to obtain both of the required holes. The parts may then not be interchangeable, depending on the care with which the holes were drilled, but the holes will be aligned and the fastener or cable will pass through properly. Similarly, it may be better to drill a passageway hole after parts are assembled. Then the hole is guaranteed to pass smoothly through all of the required parts.

*Pieces of wood may be joined together* in several different ways, the quickest being staples or nails. Both nails and staples can produce strong joints, but they can also easily split the wood if they are improperly sized or used. Wood pieces can also be glued together using white glue or carpenter's glue, but the glue joints must be made along the grain on both pieces. Pieces of wood cannot be successfully glued together if their glue is applied across end grains. The need to glue along the grain is one of the reasons that the *mortise-and-tenon* joint is used in to make wood cabinets and furniture. The strongest wood joint is formed when gluing is combined with a fastener such as a nail.

A wood screw is used to join two pieces of wood together or to fasten another material, such as plastic, to a wooden surface. The top piece of wood should have a clearance hole drilled so that the screw can slide through without touching the hole walls. The bottom piece should have a pilot hole drilled to keep the wood from splitting. Table A.1 gives approximate dimensions for both clearance and pilot holes for different screw sizes. The actual size will depend on the hardness and moisture content of the wood.

Rub a screw with soap or wax before screwing it into hardwood. If brass screws are wanted, thread their holes with a steel screw of the same size before screwing in the brass screw. Brass is much softer than steel and the screw might be damaged (especially in hardwood) if it is used both to thread the hole and to hold the pieces together.

The heads of wood screws are flat, oval, or round. Both oval and round heads will protrude above the wood's surface of the wood. The clearance hole for either a flat head or an oval head screw should be countersunk using a *countersink* to make that small conical depression. The good news is that a properly done wood screw joint will be much stronger than a nail joint and run almost no risk of splitting the wood. The bad news is that wood screws are more expensive than nails and drilling clearance and pilot holes takes some time.

*Pieces of metal may be joined* by several methods. One is by drilling clearance holes through both pieces, passing a bolt through the holes and fastening it with a nut on the far end. The clearance hole on the top piece should be countersunk if either a flat or oval head screw is used. Another approach would be to drill a clearance hole in the first piece and then drill and tap a hole in the second piece. A round head machine screw is then passed through the first piece and screwed into the second. The clearance hole on the top piece can be *counter-bored* (drilled with a flat-bottomed hole slightly larger than the screw head) if we don't want the screw to protrude above the top of the piece. It is bad practice to countersink the top piece if the bottom piece is threaded, because the screw head may snap off as it is tightened. In the parlance of geometric dimensioning and tolerancing (GD&T) (see Appendix B for geometric dimensioning and tolerancing), we have a fixed–fixed fastener and any error in position will result in huge shear stresses on the fastener.

We *assemble* our prototype once we have fabricated or bought all of the parts we need. Our choice of assembly tools will depend on our choices of materials and fasteners. A hammer is useful for nails and pins, and for shaping malleable metals. It is also useful for

**TABLE A.1 Dimensions of screws and clearance and pilot holes for steel screws used to join two pieces of wood**

| | in. | | | | mm | | |
|---|---|---|---|---|---|---|---|
| Screw Size | Diameter | Clearance Hole | Pilot Hole (Softwood) | Pilot Hole (Hardwood) | Clearance Hole | Pilot Hole (Softwood) | Pilot Hole (Hardwood) |
| 0 | 0.060 | 1/16 | None | 1/32 | 1.6 | None | 0.8 |
| 1 | 0.073 | 5/64 | None | 1/32 | 2 | None | 0.8 |
| 2 | 0.086 | 3/32 | None | 3/64 | 2.4 | None | 1.2 |
| 3 | 0.099 | 7/64 | None | 1/16 | 2.8 | None | 1.6 |
| 4 | 0.112 | 7/64 | None | 1/16 | 2.8 | None | 1.6 |
| 5 | 0.125 | 1/8 | None | 5/64 | 3.2 | None | 2 |
| 6 | 0.138 | 9/64 | 1/16 | 5/64 | 3.6 | 1.6 | 2 |
| 7 | 0.151 | 5/32 | 1/16 | 3/32 | 4 | 1.6 | 2.4 |
| 8 | 0.164 | 11/64 | 5/64 | 3/32 | 4.5 | 2 | 2.4 |
| 9 | 0.177 | 3/16 | 5/64 | 7/64 | 5 | 2 | 2.8 |
| 10 | 0.190 | 3/16 | 3/32 | 7/64 | 5 | 2.4 | 2.8 |
| 11 | 0.203 | 13/64 | 3/32 | 1/8 | 5.5 | 2.4 | 3.2 |
| 12 | 0.216 | 7/32 | 7/64 | 1/8 | 5.5 | 2.8 | 3.2 |
| 14 | 0.242 | 1/4 | 7/64 | 9/64 | 6.5 | 2.8 | 3.6 |
| 16 | 0.268 | 17/64 | 9/64 | 5/32 | 7 | 3.6 | 4 |
| 18 | 0.294 | 19/64 | 9/64 | 3/16 | 7.2 | 3.6 | 5 |
| 20 | 0.320 | 21/64 | 11/64 | 13/64 | 8.5 | 4.5 | 5.5 |
| 24 | 0.372 | 3/8 | 3/16 | 7/32 | 9 | 5 | 5.5 |

tapping close-fitting parts together. Screwdrivers should match the types of slots on your screws and bolts. Wrenches should be used for nuts and bolts, and life is much easier if at least one adjustable end wrench is at hand. Pliers are used to hold and squeeze things, but should never be used to hold a nut or a bolt. Instead, find and use a wrench that fits properly. If a model is held together with screws, a power screwdriver can sharply reduce hand fatigue and aching.

Finally, try to find a large work surface and keep it clean and orderly as assembly progresses. If a process router or assembly sequence was developed during planning, try to follow it as far as possible. Also, it is best to dry-fit parts in place before gluing or fastening them to make sure that they actually do go together properly. One of the most common mistakes in assembling prototypes is to forget that you need access to a fastener to put it in and tighten it, so put parts on the inside of an enclosed space before enclosing that space.

## A.4 SELECTING A FASTENER

A crucial aspect of almost all objects or devices that have more than one part is the nature of *fasteners* that are used to join the device's parts to each other. Fasteners and fastening methods are categorized as *permanent*, meaning that the fastener cannot be undone, or *temporary*, meaning that the fastener can be undone in a nondestructive manner. Welds, rivets, and some adhesives are instances of permanent fasteners. Screws, nuts and bolts, and paper clips are examples of temporary fasteners. There are tens of thousands of different fasteners. For example, a quick search of one distributor's website showed that we could order 78 different sizes of zinc-plated, steel Phillips flat head wood screw. Since it would be impossible to cover all existing fasteners—why would we select a Truss Opsit® Self-Tapping Left-Handed Thread Screw? We will describe just the most common fasteners and the reasons for selecting them.

Fastener selection is typically done during both the preliminary and detailed design stages. It is worth noting that each fastener is designed to meet some objective(s), satisfy some constraint(s), and serve some function(s). Thus, in addition to being of practical importance in design and model fabrication, fastener selection represents an implementation of basic design concepts. We will organize our discussion of fasteners first by material (e.g., wood, plastics, and metals) and secondarily along the distinction between permanent and temporary fasteners.

### A.4.1 Fastening Wood

Wood fastening or joining is usually performed with adhesives such as white glue, impact fasteners such as nails and staples, wood screws, or craft joints such as dovetails or dowel pins. Most wood adhesives and impact fasteners are permanent fasteners. Wood screws are normally temporary fasteners. Craft joints are usually permanent, but can be temporary. Proper craft joints usually involve a fair amount of expertise in woodworking, so we will not cover them here. There is a lot of information available on the Web about woodworking generally and about dovetail joints, mortise-and-tenon joints, and much, much more.

***A.4.1.1 Permanent wood fasteners*** We will limit our discussion to the most common adhesives for wood joining: white glue, carpenter's glue, hot-melt glue, contact cement, and nails.

*White glue* is inexpensive and strong if used properly. It is not moisture or heat resistant, so is not appropriate for use outdoors or in high-temperature environments. The fumes are not hazardous.

*Carpenter's glue* has close to the same consistency as white glue. It is strong if used properly, and has moderate moisture and heat resistance. It fills gaps well. The fumes are not hazardous.

*Hot-melt glue*, often just called *hot glue*, melts at high temperature and solidifies at room temperature. It is applied with a glue gun that heats the glue to its melting point. Its strength is moderate to low, and it is moisture-resistant but not heat-resistant. It is excellent for quick assembly and short-lived prototypes. It is easy to burn yourself with the hot glue, so use caution with it.

*Contact cement* is most often used to bond veneers or plastic laminates to wood. It is extremely permanent, strong, and heat- and moisture-resistant. It must be applied to both surfaces to be joined and allowed to dry until tacky. Then the two surfaces are joined. The fumes are hazardous, as is contact with the uncured cement, so protective gloves, goggles, and adequate ventilation are needed.

*Nails* are considered permanent fasteners even though they can sometimes be removed without permanent damage—we should not count on always being able to remove a nail. A nail holds two boards together by friction and by the nail head, if there is one. Nails are not normally considered precision fasteners and will display a fairly wide range of dimensions as manufactured. Practically, a nail should be sized so that approximately 2/3 of its length is in the bottom board.

There are many different kinds of nails, of which *common nails* are the most common (see Figure A.1). They are sized in *pennies* (the approximate weight in pounds of 1000 nails), which is abbreviated as "d." They range from 2d, which are 1 in. long and made from 15-gage wire, to 60d, which are 6 in. long and made from 2-gage wire. A common nail is for general-purpose joining of boards.

*Finishing nails* have a small, almost nonexistent, head (Figure A.1), and are slightly smaller in diameter than common nails. Finishing nails are countersunk with a nailset so their heads are below the wood's surface. They are used for cabinetry and in other circumstances where the nail head should not show.

*Box nails* (Figure A.1) are used to join thin pieces of dry wood. They have a blunt tip to avoid splitting the wood. Box nails also have slightly smaller diameters than common nails, and they often have a coating that heats and melts as they are driven: The coating then solidifies and glues the nail in place.

*Brads* are small wire nails that resemble small finishing nails (Figure A.1). They are typically used for attaching molding to walls or other places where small inconspicuous nails are required.

***A.4.1.2 Temporary wood fasteners*** Wood screws are the most common temporary fasteners in wood. Wood screws come with flat heads, oval heads, or round heads. They are typically made of three materials: brass, galvanized steel, or stainless steel. Brass is generally used for decorative applications as it is soft and easily damaged. Stainless steel is the most expensive, but it is the most resistant to rust and corrosion. Galvanized steel is the most common. Table A.2 lists standard wood screw sizes, their corresponding diameters, and their clearance and pilot hole sizes. Wood screws range in length from $^1/_2$ in. to $3^1/_2$ in.

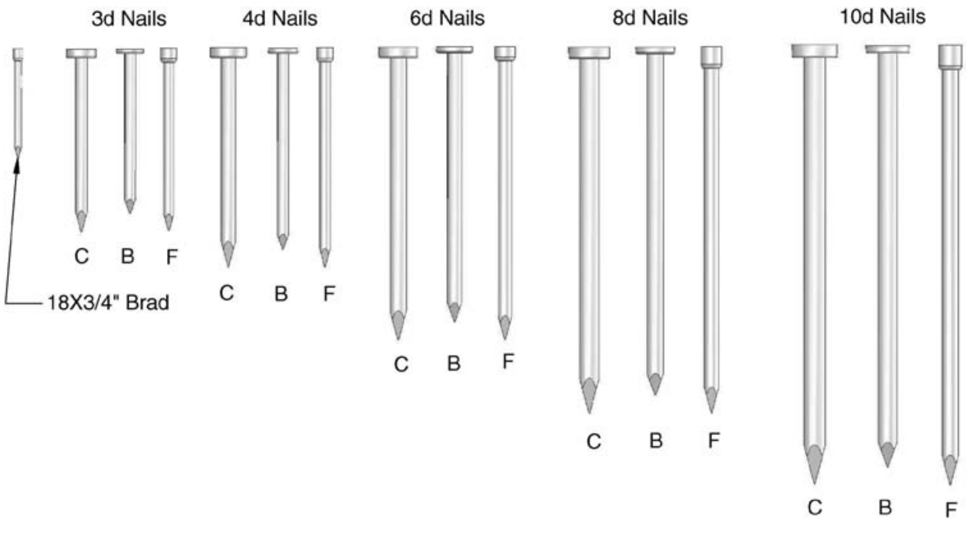



**FIGURE A.1** A collection of four styles of standard nails of varying size: a brad, common nails (C), box nails (B), and finishing nails (F). Courtesy of R. Erik Spjut.

**TABLE A.2 Common *inch* screw sizes and dimensions and clearance hole dimensions for machine screws**

| Screw Size | Major Diameter | Pitch Diameter | Minor Diameter | Normal Clearance Hole | Close Clearance Hole |
|---|---|---|---|---|---|
| 0-80 | 0.060 | 0.052 | 0.044 | 0.073 | 0.067 |
| 1-64 | 0.073 | 0.063 | 0.053 | 0.089 | 0.081 |
| 2-56 | 0.086 | 0.074 | 0.064 | 0.106 | 0.094 |
| 3-48 | 0.099 | 0.086 | 0.073 | 0.120 | 0.106 |
| 4-40 | 0.112 | 0.096 | 0.081 | 0.136 | 0.125 |
| 5-40 | 0.125 | 0.109 | 0.094 | 0.154 | 0.140 |
| 6-32 | 0.138 | 0.118 | 0.099 | 0.169 | 0.154 |
| 8-32 | 0.164 | 0.144 | 0.125 | 0.193 | 0.180 |
| 10-24 | 0.190 | 0.163 | 0.138 | 0.221 | 0.205 |
| 1/4-20 | 0.250 | 0.218 | 0.188 | 0.281 | 0.266 |
| 5/16-18 | 0.313 | 0.276 | 0.243 | 0.344 | 0.328 |
| 3/8-16 | 0.375 | 0.334 | 0.297 | 0.406 | 0.390 |
| 7/16-14 | 0.438 | 0.391 | 0.349 | 0.469 | 0.453 |
| 1/2-13 | 0.500 | 0.450 | 0.404 | 0.531 | 0.516 |

*Round head screws* (Figure A.2) protrude above the surface of the wood (typically for cosmetic reasons) and the screw head rests flush against the top surface of the wood. They are most often used for mounting hardware such as hinges or knobs onto wood.

*Oval head screws* (Figure A.2) resemble a cross between a flat head and a round head screw. The head is designed to protrude above the surface of the wood (again for cosmetic

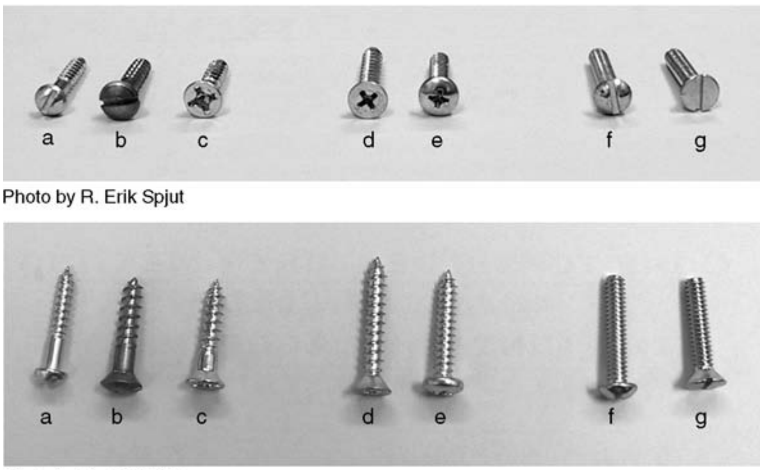



Photo by R. Erik Spjut

Photo by R. Erik Spjut

**FIGURE A.2** Assorted screws (temporary fasteners), each a No. 10 size, and all but (d) 1 in. long: (a) steel slotted round head wood screw; (b) brass slotted oval head wood screw; (c) steel Philips flat head wood screw; (d) steel Philips flat head sheet metal screw ($1\frac{1}{4}$ in. long); (e) steel Philips pan head sheet metal screw; (f) steel slotted round head machine screw; and (g) steel slotted flat head machine screw.

reasons), but the hole should be countersunk. Oval head screws are most often used to attach pre-countersunk hardware, such as hinges, to wood.

*Flat head screws* (Figure A.2) are used where the screw cannot protrude above the surface of the finished wood. The hole for the screw should be countersunk unless the wood is particularly soft, in which case the screw can simply be driven so that its head is below the surface of the wood.

Screw slots come in slotted, Philips, and specialty varieties. The specialty slot, such as a Torx, requires a special head to drive it, and is used in such applications as closing the clamshell of the DeWalt D21008K corded power drill pictured in Figure A.3. Fasteners with Philips heads can support greater driving forces than those with slotted heads; they are preferred if the screws are to experience high torque while being tightened. Some of the specialty slots can support even greater driving force than the Philips type.

## A.4.2 Fastening Polymers

Permanent polymer joining is most commonly done with adhesives that generally fall into two classes, solvent-based cements that are polymer specific and general adhesives such as epoxy. The general adhesives must generally be used when joining different polymers or a plastic to wood or metal. In rare cases, polymers are joined using friction welding. Temporary fasteners are usually threaded fasteners such as machine screws and nuts and bolts, very much like those used for metal, except that the fasteners may be made from a polymeric material such as nylon or acetyl resin. Accordingly, our discussion of threaded fasteners will be deferred to the section on fastening metals.

***A.4.2.1 Permanent polymer fastening*** When they can be used, *solvent cements* are the preferred means of joining two pieces of a polymer together. If done properly, the joint will have the same strength and characteristics as the bulk material. Typically, the solvent

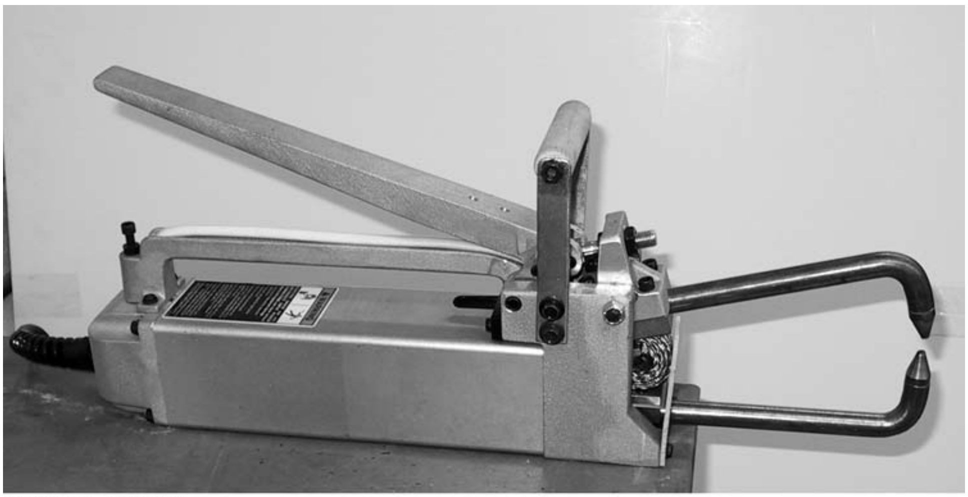


Photo by R. Erik Spjut

**FIGURE A.3** A Miller LMSW-52 Spot Welder.

will dissolve some of the original material and then evaporate, permitting the material to resolidify. Solvent cements work best when the two surfaces being joined have a nearly perfect physical joint with no gaps or holes. (Some thicker solvent cements can fill gaps.) If too much solvent cement is used, the material may be weakened. Specific solvent cements include plastic model cement for polystyrene, primer and cement for PVC piping, and acrylic solvent cement for joining acrylic. Each of these solvent cements typically comes with its own instructions that should be carefully read and followed exactly.

*General adhesives* should be chosen after it has been determined whether or not a particular adhesive is recommended for joining the chosen materials. The first to be examined should be epoxies and cyanoacrylates (superglues). Epoxies do very well in bonding porous materials and do a good-to-poor job of bonding nonporous material, depending on the specific material. Cyanoacrylates work very well on smooth nonporous materials but do poorly on porous materials. Contact cements should be examined next if epoxies or cyanoacrylates prove unacceptable.

***A.4.2.2 Temporary polymer fastening*** Our discussion of threaded plastic fasteners is deferred to the corresponding section on metal fastening because they are similar to temporary metal fasteners.

## A.4.3 Fastening Metals

The principal permanent means of joining metals are soldering/brazing, welding, and riveting. Threaded fasteners are the principal temporary means of joining metals.

***A.4.3.1 Fastening or joining metals together permanently*** *Welding* involves melting portions of the two pieces to be joined and (usually) adding some additional metal. The joint is formed when the metal resolidifies. Welding is most often used to join two pieces of ferrous metals (steels and cast iron), but it can be done on aluminum and other metals, by expert welders under the correct circumstances. Arc welding involves specialized training and equipment and is well beyond the level of this text.

*Spot welding* is done to join two pieces of (usually ferrous) sheet metal, usually with a relatively low-cost and safe *spot welder* that consists of two long arms that end in electrode tips (see Figure A.3). The two pieces of sheet steel are squeezed between the two electrode tips and a brief-but-large current is passed through the electrodes and the sheet steel. The current resistively melts a small spot between the sheets that then solidifies and forms the joint. (As do other power tools, the spot welder has its own procedural and safety instructions that should be attentively followed.)

*Soldering* and *brazing* join two pieces of higher-melting-temperature metal with a piece of low-melting-temperature metal. The difference between the two is the temperature at which the joining metal melts. By convention, using a joining metal that melts below 800 °F (425 °C) or 450 °C (840 °F) is soldering, and using a joining metal that melts above that is brazing. Depending on the size of the joint and the temperature involved, a soldering iron, a soldering gun, or a butane torch can be used for the heating and melting. It is important to have good mechanical contact between the pieces to be welded or soldered before heating. The molten metal will be drawn into the gap between the pieces by capillary action. The joint is not as strong as a weld, but can be done with much less training or specialized equipment.

*Rivets* are the final common permanent fastener used to join metal parts. There are two principal types: *Solid rivets* are used when there is access to both sides of the joint and maximum strength is required. *Blind rivets*—often called pop rivets, although POP® is a registered brand name for a blind rivet—which can be installed when there is access only to one side of the joint. Solid rivets require specialized training and equipment and are mentioned only in passing.

Blind rivets are installed with a rivet gun. A hole is drilled through the two pieces to be joined, in accord with the manufacturer's recommendation. A blind rivet is placed in the rivet gun, and then inserted into the hole. The handle on the rivet gun is squeezed until the mandrel snaps off. Figure A.4 shows such a rivet gun and the procedure for installing a blind rivet.

Rivets can support tensile loads but are more often used for shear loads. The machine screws discussed in the next section are used for tensile loads.

***A.4.3.2 Fastening metals together temporarily*** The principal temporary fasteners for metal are sheet metal screws, machine screws, cap screws, and bolts and nuts. There is no universal definition that differentiates a screw from a bolt. Some maintain that screws are threaded fasteners that come to a point and bolts are threaded fasteners

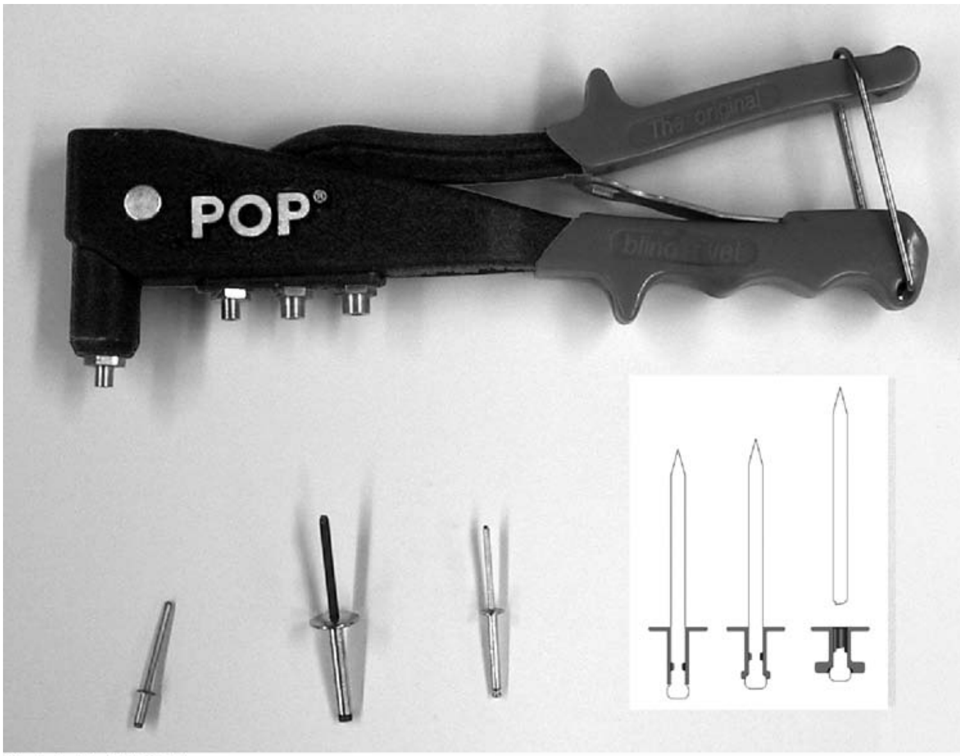



Photo by R. Erik Spjut

**FIGURE A.4** A Pop® pop rivet gun (top) and a graphic showing just how a rivet is actually installed (bottom).

with a constant thread diameter and a square end. By this definition, machine screws are actually bolts. Others argue that screws are meant to be turned or rotated while being attached, and bolts are designed not to rotate during or after attachment and that they (usually) have smooth non-grippable tops. Under this definition, a hex head bolt is a screw. The lesson for us is that we must exercise care when we refer to screws or bolts; perhaps it's best to follow local custom. An assortment of threaded fasteners is shown in Figures A.2 and A.5.

*Machine screws* come in a wide variety of head types, slot types, materials, diameters, and lengths. The most common head types are pan, round, cheese, flat, and oval. There are many variants on these basic types. The choice of head type depends on whether the resultant joined surface must be flush and whether the fastener can be fixed or floating. Fixed and floating fasteners will be discussed in Appendix B, but briefly: The location of a fixed fastener cannot be adjusted in position as it is tightened, and the location of a floating fastener can be adjusted slightly as it is tightened. Flat heads require countersinking of the surface and result in a flush surface, but a fixed fastener (Figure A.2(g)). Oval heads require countersinking and result in a rounded but protruding surface, and a fixed fastener. Pan, round, and cheese heads (and no, not the Green Bay type) all result in protruding heads, but floating fasteners. The difference is the degree to which they protrude. However, if there is room, the holes for all three

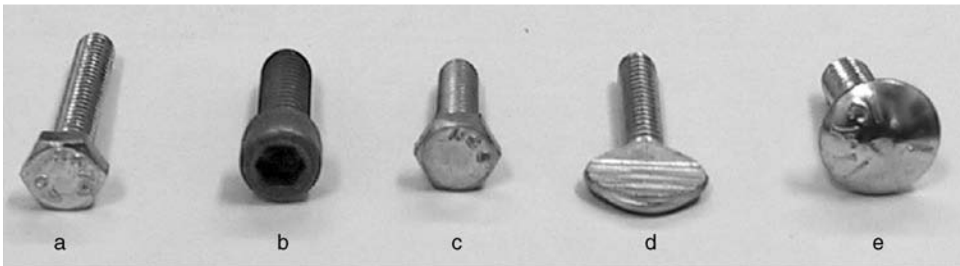



Photo by R. Erik Spjut

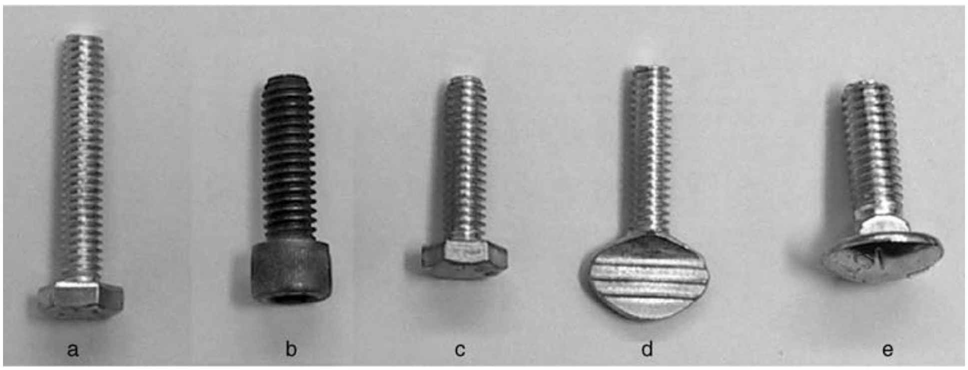



Photo by R. Erik Spjut

**FIGURE A.5** Assorted screw and bolt (temporary) fasteners: (a) steel hex bolt 1/4-20 × 2 in.; (b) black oxide steel socket head cap screw 5/16-18 × 1 in.; (c) steel hex cap screw 1/4-20 × 1 in.; (d) steel thumb screw 1/4-20 × 1 in.; and (e) steel carriage screw 5/16-18 × 1 in.

may be counter-bored, resulting in a flush surface, but with a noticeable gap around the screw head (Figure A.2(f)).

*Cap screws* are sometimes considered machine screws and sometimes considered as a separate category. Cap screws have either hex heads or socket heads. The hex heads are designed to be tightened with a wrench. Hex head cap screws are almost never used in counter-bored holes due to the difficulty of getting a wrench in the hole to tighten the screw (Figure A.5(c)). Socket head cap screws are designed to be tightened with a hex key or Allen wrench and are frequently used in counter-bored holes to leave a flush surface (Figure A.5(b)).

Machine screws are most commonly made of steel, stainless steel, aluminum, brass, or nylon. There are others designed for specific applications. The material chosen is a function of cost, strength needed, and compatibility with the metals being joined.

The dimensions of machine screws are governed by standards. *Inch fasteners* are specified with a thread diameter and the number of threads per inch (TPI). Diameters smaller than $^1/_4$ in. are specified with a gauge number. A $^1/_4$-20 is an inch fastener with a $^1/_4$ in. thread diameter and 20 TPI. *Metric fasteners* are specified with a thread diameter and a thread pitch (the distance between adjacent threads). An M6 $\times$ 1 is a metric thread fastener with a 6 mm thread diameter and a thread pitch of 1 mm. Common *inch* screw sizes and clearance hole dimensions are listed in Table A.2, and common *metric* thread sizes and clearance hole dimensions are listed in Table A.3.

When we specify clearance holes for threaded fasteners, we must take into account the skill of the machinist and the cost of precision machining. The normal clearance holes are for reasonably competent yet inexpensive machining. The close clearance holes are for precise and more expensive machining. The available tolerance that we can specify in geometric dimensioning and tolerancing is the difference between the clearance hole and the major diameter. For example, as we will discuss in Appendix B, a $^1/_4$-20 machine screw with a close clearance hole will have only 0.266–0.250 = 0.016 in. available for tolerancing.

**TABLE A.3 Common *metric* thread sizes and dimensions and clearance hole dimensions for machine screws**

| Screw Size | Major Diameter | Pitch Diameter | Minor Diameter | Normal Clearance Hole | Close Clearance Hole |
|---|---|---|---|---|---|
| M1.6 × 0.35 | 1.60 | 1.37 | 1.17 | 1.9 | 1.75 |
| M2 × 0.4 | 2.00 | 1.74 | 1.51 | 2.5 | 2.25 |
| M2.5 × 0.45 | 2.50 | 2.21 | 1.95 | 3.0 | 2.75 |
| M3 × 0.5 | 3.00 | 2.68 | 2.39 | 3.7 | 3.3 |
| M3.5 × 0.6 | 3.50 | 3.11 | 2.76 | 4.3 | 3.9 |
| M4 × 0.7 | 4.00 | 3.55 | 3.14 | 4.8 | 4.4 |
| M5 × 0.8 | 5.00 | 4.48 | 4.02 | 5.8 | 5.4 |
| M6 × 1 | 6.00 | 5.35 | 4.77 | 6.8 | 6.4 |
| M8 × 1.25 | 8.00 | 7.19 | 6.47 | 8.8 | 8.4 |
| M10 × 1.5 | 10.00 | 9.03 | 8.16 | 11.0 | 10.5 |
| M12 × 1.75 | 12.00 | 10.86 | 9.85 | 13.0 | 12.5 |
| M14 × 2 | 14.00 | 12.70 | 11.55 | 15.0 | 14.5 |

### A.4.4 What Size Temporary Fastener Should I Choose?

It is a sad-but-true fact that the majority of threaded fasteners are chosen because they "look right" to the experienced designer. But the proper way to choose the diameter of fastener is to:

- calculate the force that the fastener is expected to endure;
- include a reasonable safety factor; and
- choose a fastener that exceeds the strength required.

The two forces that a screw is likely to experience are a tensile force (along the axis of the screw) and a shear force (across the axis of the screw). The calculation of these forces in a complex piece of machinery is beyond our scope of this book (although it can be found in a typical "strengths" or "mechomat" text). The manufacturer's specification that is of interest to us is the *proof load*, which is the load the fastener must withstand without undergoing permanent plastic deformation. We would typically choose a fastener with a proof load four times the expected maximum load (this corresponds to a safety factor $S = 4$). Also, we would normally tighten the fastener to have a *preload* of 90% of the proof load. The torque required to preload the bolt may be estimated as

$$T = 0.2 F_1 d \tag{1}$$

Here $T$ is the torque, $F_1$ the proof load, and $d$ is the nominal diameter of the fastener. For example, if the maximum tensile load is expected to be 1550 N, the proof load should be $4 \times 1550\,\text{N} = 6200\,\text{N}$. After a search we found a manufacturer that has a steel Philips pan head machine screw M6 × 1 with a proof load of 6230 N, so we would choose this screw. From Table 7.4 we note that the pitch diameter is $5.35\,\text{mm} = 0.00535\,\text{m}$. The preload is then $0.9 \times 6200\,\text{N} = 5580\,\text{N}$. We would then use a torque wrench to tighten this screw to a torque of $0.2 \times 5580\text{N} \times 0.00535\,\text{m} = 29.9\,\text{N m}$.

In closing, we would note that the subject of fastener selection has filled many a volume and more than a few manufacturer's catalogs. The suggestion made above should be viewed as a starting point, not as the final word on fastener selection. Having said that, it is still the case that our guidelines will be adequate for designing and building most common models or prototypes. If our design is critical, or if our designer is not widely experienced with fasteners, we would seek expert guidance from a mentor, a machinist, or a reference book.

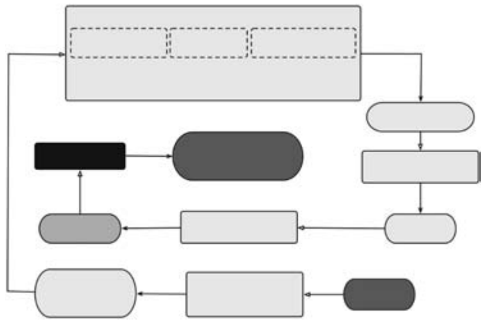
## A.5 NOTES

*Sections A.2 and A.3*: Anyone with any interest at all in woodworking should definitely read Abram (1996)! There are also many sources for data about common construction techniques and fasteners on the Internet. Particularly helpful websites are:
*Industrial Screw* <http://www.industrialscrew.com/index.cfm?page=tech>;
*Lowe's How To Library* at the Lowe's website <http://www.lowes.com>;
*Bob Vila's How To Library* <http://www.bobvila.com/HowTo_Library/>;
and *eHow* <http://www.ehow.com/>.

*Section A.4*: The ANSI B18 series covers rivets, bolts, nuts, machine and cap screws, and washers in American engineering units. The thread size is governed by the Unified Thread Standard, ANSI B1.1, ANSI B1.10 M, and ANSI B1.15. Metric screw threads are governed by ISO 68-1, ISO 261, ISO 262, and ISO 965-1. The website *This-To-That* <http://www.thistothat.com/> gives specifics on selecting adhesives for joining two materials together, that is, this to that.

# APPENDIX B

# PRACTICAL ASPECTS OF ENGINEERING DRAWING



ENGINEERS AND machinists have developed a common language for engineering drawings in order to communicate design ideas effectively and efficiently. This language is detailed in the ASME Y14.5M-1994 standard and is referred to as *geometric dimensioning and tolerancing* (GD&T). As preface to a full discussion of GD&T, we will first review best practices of dimensioning.

## B.1 DIMENSIONING

In order to understand the geometric dimensioning and tolerancing system, we must first understand the appropriate method for *dimensioning* or putting dimensions on a drawing. Dimension placement, symbols, and conventions are all important to the common language for engineers and machinists. The following concepts are essential for understanding dimensioning of technical drawings.

### B.1.1 Orthographic Views

Most technical drawings show orthographic views of the object being represented. *Orthographic* or *principal* views are drawings based on the projection of the object onto a plane. The best way to visualize an orthographic drawing is to imagine a “glass box” around the object, with a projection of the object onto each surface of the box. The box is then unfolded to give rise to the six primary views of the orthographic drawing: top, front,

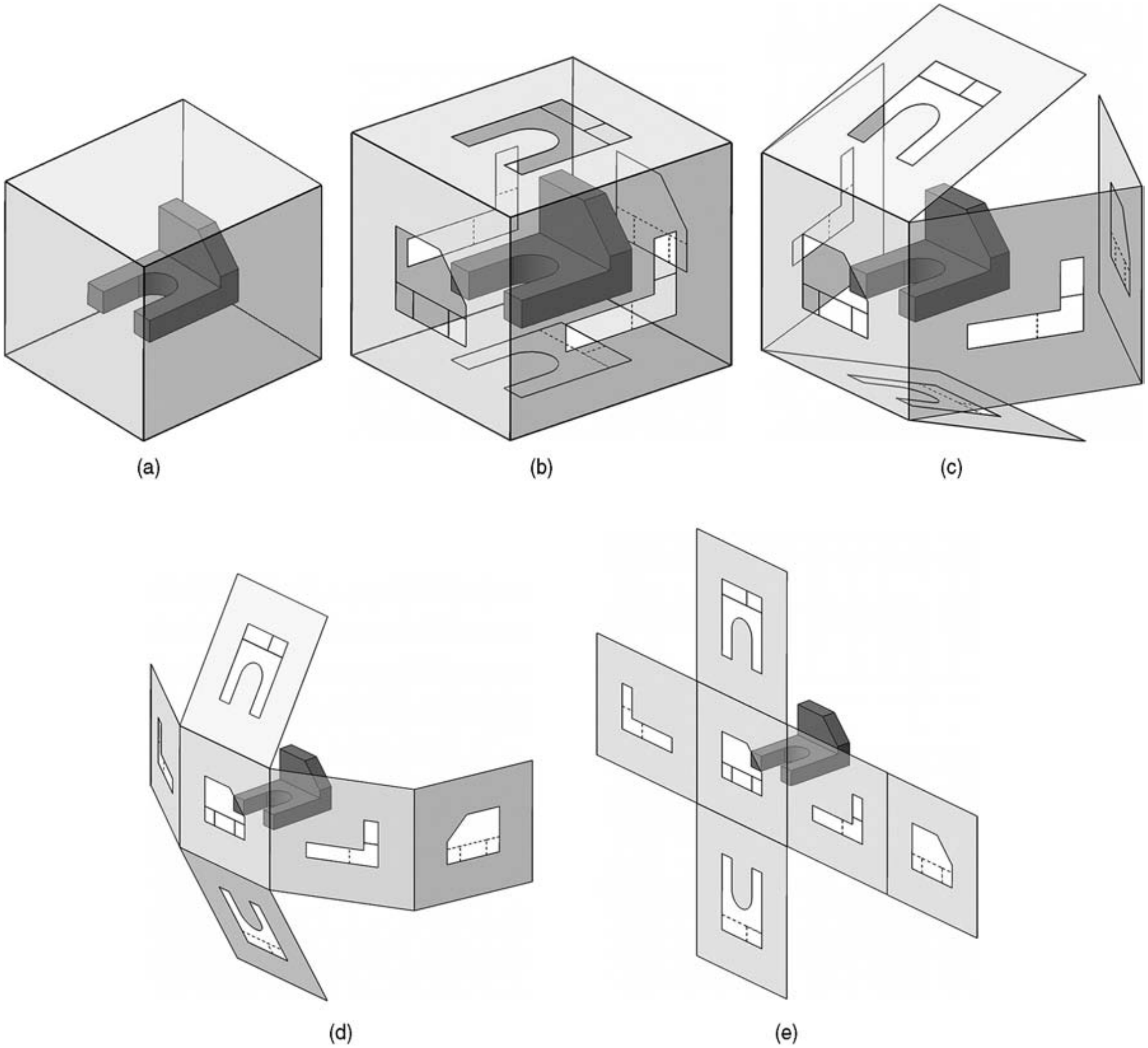



**FIGURE B.1** The six *orthographic* or *principal* views of an object. The orthographic views are created by projecting the object onto a plane. This can be visualized by imagining: (a) a "glass box" around an object; with (b) a projection of the object onto each face; and then (c–e) the unfolding of the box leads to the six views: front, top, bottom, right side, left side, and back views. Note that in practice, we often need to use only the front, top, and right side views to fully describe the object as the others are redundant. Adapted from *Engineering Graphics Essentials* by permission of K. Plantenberg.

and bottom views; and right-side, left-side, and rear views (Figure B.1). It should be noted that this particular type of orthographic drawing uses *third angle projection* in which the drawing is derived from an image projected *onto a plane between the observer and the object*, so that the order is observer, projected view, object. Or, said somewhat differently, in third angle projection the image is projected onto a plane *in front of* the object.

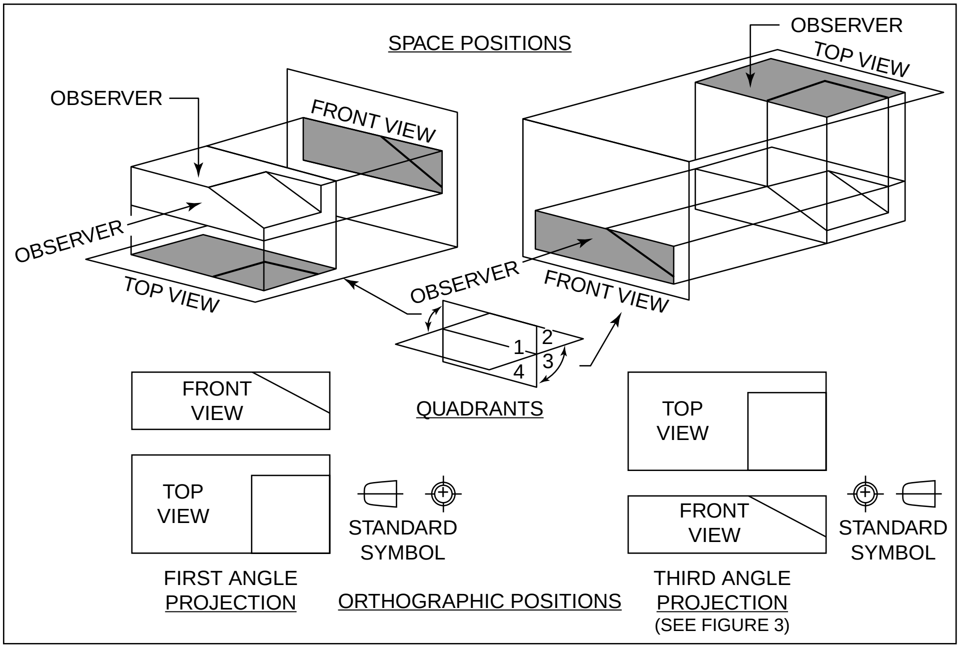



**FIGURE B.2** First and third angle projections. The difference between these two orthographic projections lies in the location of the plane on which the object is projected. In first angle projection, the object is projected onto a plane *behind* it. In third angle projection, the object is projected onto a plane *in front of* it. Note the different symbols used to represent each drawing type. Reprinted from ASME Y14.3-1975 and ASME Y14.5-1994 (R2004), by permission of The American Society of Mechanical Engineers. All rights reserved.

In Japan and some European countries, a different type of orthographic view is used: In *first angle projection* the drawing is derived from an image projected *onto a plane behind the object*, so that the order is observer, object, projected view. That is, in first angle projection the image is projected onto a plane *behind* the object. The two types of orthographic views can lead to very different drawings and it is important to know which system is being used. Figure B.2 shows the views in first and third angle projections, as well as the symbols used to denote which system is depicted. It is important to note that all six views of the orthographic projection are not always required. We can often fully define an object with front, top, and right-side views (in third angle projection), or front, bottom, and right side (in first angle projection). In some cases we need only the front and top views. It is important to note that the orthographic views are to be laid out as one drawing, that is, the three (or two) views must line up as they are laid out in the projection, with features aligning in the views.

Choosing an appropriate front view for an orthographic drawing is essential for ensuring its correct interpretation. It is much easier to figure out what is being represented given the right front view, because that front view is seen first and it represents the most basic and characteristic profile of the object being drawn. In addition, the front view should be stable (i.e., heavy on the bottom), and should have as few hidden lines

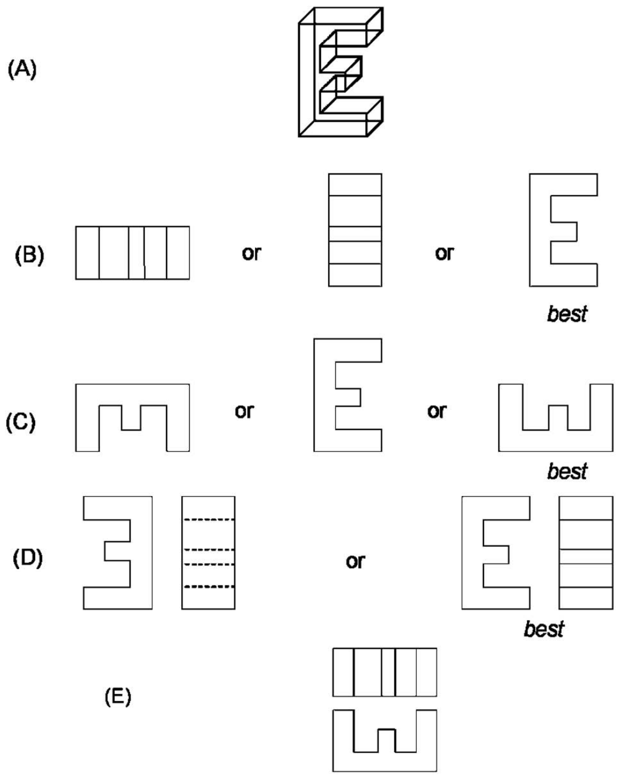



**FIGURE B.3** Choosing a front view: (a) isometric view of object to be drawn; (b) front view should be chosen to show the most informative profile of the object; (c) the front view should be chosen so that the view shows the most stable version of the object; (d) a front view should be chosen to minimize the hidden lines of an object in other views; and (e) the best choice of views for this object: front and top views.

as possible. Consider a block letter "E," as in Figure B.3. Several poor choices for front views of this object are shown, as well as several "best" front views.

### B.1.2 Metric versus Inch Dimensioning

Metric and inch dimensions (and that *is* what they're called) are specified differently in the ASME standard, which enables us to tell at first glance which system of units is used on a drawing. American (inch or in.) dimensions are specified to have no zero before the decimal point (e.g., .5 in.). Further, the dimension must contain the same number of decimal places as the tolerance for that dimension. For example, if the tolerance for a given dimension is .01 in., the dimension must be .50 in. Metric dimensions include a zero before the decimal point (e.g., 0.5 m). A metric dimension does not need to match the number of decimal places with the tolerance, and no decimal point or zero is included if the dimension is a whole number.

### B.1.3 Line Types

Technical drawings use several different types of lines. The weight and style of these lines, as well as their placement on the drawings, are specified in the ASME standard. Most CADD packages include settings for the ASME standard, so that they will automatically place the lines correctly if they are initially set correctly. *Extension lines* come out from a part and leave a visible gap between the part and the line. Examples of extension lines can be seen in Figure B.4; these are the vertical lines up from the screwdriver handle. *Dimension lines* are typically broken for the numbers and are placed at least 10 mm apart from the object on the drawing. Subsequent dimension lines are at least 6 mm apart. Examples of these lines can also be seen in Figure B.5. *Leader lines* are used to indicate surfaces and hole diameters. They should be at an angle between 30 and 60°, point toward the center of a hole, and include only one dimension per leader line. The left-side view of the screwdriver print shows a leader line pointing to the outer diameter of the piece; note that the arrow points directly at the center of the diameter. *Hidden lines* are dashed lines (- - - - - - - -) that indicate the presence of a feature that is seen in another view. Hidden

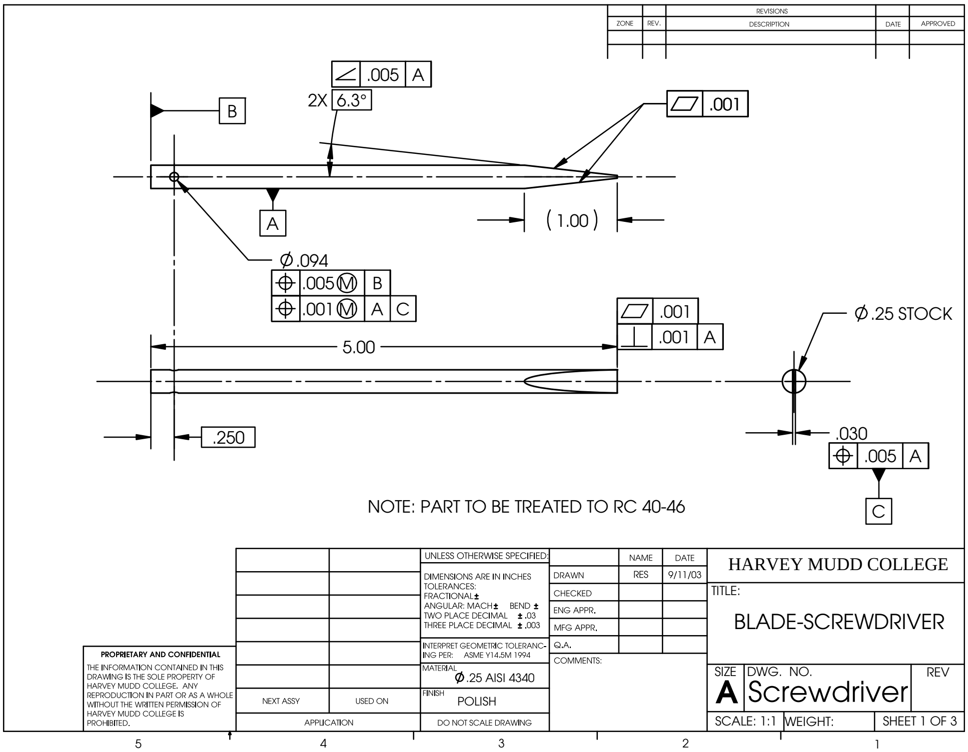



**FIGURE B.4** A drawing of a screwdriver blade indicates all the types of dimensions: *basic* (indicated by boxing of the numbers); *reference* (indicated by parentheses); *stock* (indicated by STOCK following the dimension); and *size/location dimensions* (for example, the overall blade length, 5.00 in.). Courtesy of R. Erik Spjut.

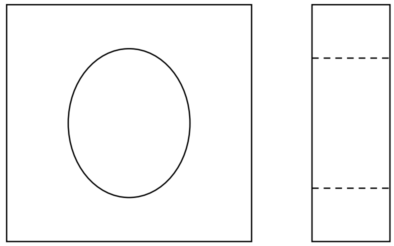



**FIGURE B.5** Hidden lines are indicated by dashed lines. They are used to represent a feature in a view where that feature is not explicitly seen. For example, the hole shown in the front view is located by hidden lines in the side view.

lines indicate the presence of a hole in the object shown in Figure B.4. The front view shows the hole; the side view uses hidden lines to indicate where the hole is located. *Center lines* are used to indicate a cylinder and are represented by the type of dashed line shown in Figure B.5. The presence of this type of line alone indicates a cylindrical feature; the side view showing the part is a cylinder is not needed.

## B.1.4 Orienting, Spacing, and Placing Dimensions

The most common practice is to orient all dimensions such that they can be read when the drawing is held horizontally. It is also acceptable to use an aligned system in which dimensions are oriented either vertically such that they can be read from the right or horizontally such that they can be read from the bottom. As mentioned before, the minimum spacing between adjacent dimensions is specified as 6 mm. The placement of dimensions is also important. Dimensions should be stacked with the shortest dimensions placed closest to the object and longer dimensions beyond them. This system avoids crossing extension lines, thereby minimizing confusion. Dimensions should also be staggered to make it easier to read them. In addition, in orthographic drawing views, the dimensions should be placed *between* the drawings views, that is, between the front and top views, and between the front and right-side views. The overall size, length, height, and depth should be specified.

## B.1.5 Types of Dimensions

It is important to distinguish between size dimensions and location dimensions. *Size dimensions* define the size of features: overall height, length, thickness, diameter of a hole, size of a slot, and so on. *Location dimensions* specify where a feature is located with respect to other features or the edge of an object. Location dimensions define the center of a hole or the location of a slot, for example, with respect to the edge of a part or with respect to another feature. The general rule of thumb is to dimension size dimensions first, then do location dimensions. Remember that both size and location dimensions will have tolerances associated with them, a concept that we will revisit later in this chapter.

In addition to size and location dimensions, there are three other dimension types that are important: basic dimensions, reference dimensions, and stock dimensions. Figure B.6 is a drawing of a screwdriver blade, from the same screwdriver whose handle appears in

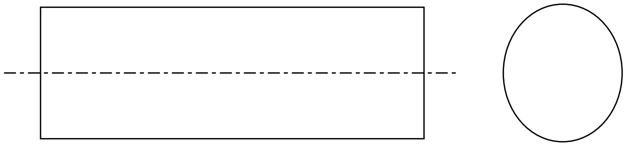



**FIGURE B.6** Center lines are indicated by a different type of dashed line and describe cylindrical features.

Figure B.5, and it illustrates all of these types of dimensions. (And by the way, are the dimensions in this drawing in millimeters or inches?) First, the overall blade length in Figure B.6 is an example of the *size* dimensions above. The boxed numbers in the drawing are *basic dimensions* (e.g., the .250 dimension from the blade end to the hole on the left side of the drawing). Basic dimensions define the basis for permissible variation, or tolerance, in the geometric dimensioning and tolerancing system. In other words, they define the theoretically exact point from the end of the blade from which to measure the variation in the location of the hole. We will revisit this concept in the next section on tolerancing. *Reference dimensions* are indicated in parentheses, for example, the (1.00) dimension on the screwdriver blade in the top view. A reference dimension is a point of information for the machinist and is not a requirement. It means that if the part has been produced correctly, the blade length should be around 1 in. The last type of dimension is a *stock dimension*, and it is indicated by writing .25 STOCK on the drawing. This indicates that the material used for the part comes from the manufacturer with a specified size and associated tolerance; no further tolerance specification is required.

*Every* specified dimension requires an associated tolerance in a technical drawing except basic, reference, and stock dimensions. This makes sense when we realize the purpose of these types of dimensions. One more important note should be made: Any number on a technical drawing that does not have a tolerance directly associated with it (e.g., $.50 \pm .01$), still has a specified tolerance on the drawing. These tolerances are specified in the title block of a technical drawing and are called *block tolerances*. In Figure B.6 the block tolerances can be seen to be $.XX \pm .03$. The tolerance is determined by the number of decimal places in the dimension. This means that the 5.00 overall blade length has a tolerance of $\pm.03$ in.

## B.1.6 Some Best Practices of Dimensioning

Each of the concepts and dimension types described above can be integrated into a set of guidelines for dimensioning a technical drawing. Three of the most important rules for dimensioning are as follows:

- Dimension the *size dimensions first*, then do the location dimensions.
- *Do not double dimension*. In an orthographic drawing, it is not necessary to specify the same dimension twice. For example, specifying the depth of an object on the right and top views leads to unnecessary clutter in the drawing. It is not good drawing practice.

- *Do not dimension to hidden lines*. Dimension a feature where it is visible. A hole, for example, should be dimensioned in the view where it is visible. This is good practice that leads to clearer technical drawings.

## B.2 GEOMETRIC TOLERANCING

Now that we have covered some basic dimensioning symbols and rules, we turn to geometric tolerancing. A *tolerance* is the permissible variation of a part. Tolerances are applied to all size and location dimensions. Tolerances are required because we, as engineers, need to know how much a part can vary from its specifications before it no longer functions as intended. Defining tolerances requires that we know and understand the function of a given part. It is good practice to specify tolerances only as tightly as we need because parts become much more expensive to manufacture as their tolerances become smaller. Figure B.7 shows the relative cost of increased tolerances. Here the $y$-axis tracks the percentage of increase in cost of making the hole, and the $x$-axis shows both the size and location tolerances on the hole. (Location tolerances will be described below.) This figure not only gives us an idea of the increased cost as the tolerances become tighter, but also tells

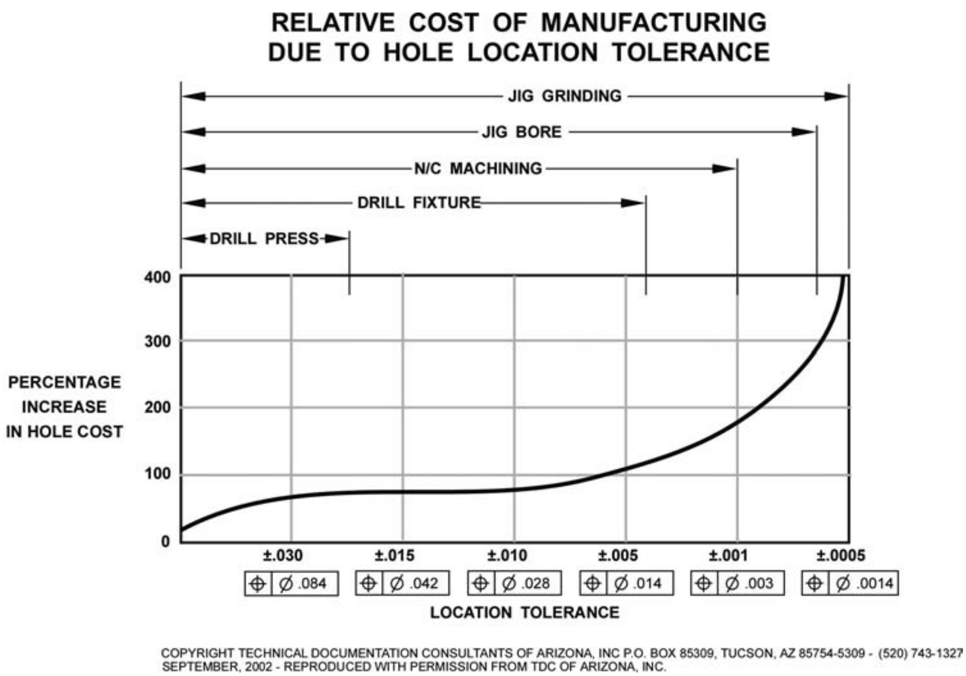



**FIGURE B.7** Relative cost of manufacturing as hole location tolerance gets smaller. The cost goes up significantly as smaller tolerances are prescribed. In addition, special equipment is required to meet tight tolerances. Reprinted by permission of Technical Documentation Consultants of Arizona, Inc.

us what type of machinery is required to make such a hole. It is not surprising that the more precise a hole it can make, the more expensive the equipment!

All dimensions require a tolerance (except for the basic, reference, or stock dimensions described above). On a drawing, there are several places to look for tolerance specifications:

- associated with a dimension (+/−);
- in a feature control frame (which we describe below);
- in a drawing note; or
- in the block tolerance as the default if no other tolerance is applied.

It is possible to tolerance every dimension with a plus or minus tolerance, but the geometric dimensioning and tolerancing system provides more leeway for each part, which in turn leads to cost savings. The geometric tolerancing system takes into account not only variations in *size* of an object, but also permissible variations on the *position*, *form*, and *orientation* of features. We now describe some of the tolerancing components of the GD&T system.

### B.2.1 The 14 Geometric Tolerances

There are 14 characteristics specified in the ASME Y14.5M-1994 standard that can vary, and therefore have an associated tolerance (Figure B.8). For example, we can specify how much a surface can vary in flatness or how much variation is permissible in the location of a hole. These 14 characteristics are categorized into five groups: form, profile, orientation, location, and runout. These groups are somewhat hierarchical. For example, a position tolerance is a refinement of an orientation tolerance, which is a refinement of a form tolerance, which is a refinement of the size tolerance on a feature. For example, if a rectangular piece is .500 ± .004 in. in height, the minimum height is .496 and the maximum is .504, simply based upon the size dimensions. If each end of the part was made at one of these extremes—the part would be within the size tolerance—the maximum out-of-flatness that the top surface could be is .008. Therefore, if a flatness tolerance is to be applied to this part, it must be *less than* .008 in. for it to make sense.

Form tolerances apply to individual features, for example, a surface in the case of straightness or flatness. All other tolerances apply to related features. For example, orientation and location tolerances specify permissible variation of a given feature with respect to a reference frame. Therefore, these tolerances require specification of a reference frame in order for them to be meaningful. The reference frames are defined by datums, which we will soon discuss below.

A full discussion of all of the 14 geometric tolerances is beyond our scope (see the notes in Section .4 for further reading). Therefore, we will focus specifically on position tolerances to show how geometric tolerances are applied.

### B.2.2 Feature Control Frames

Feature control frames are devices used to specify the particular geometric tolerance on the technical drawing. We have seen them in the drawings presented earlier. The feature

| | TYPE OF TOLERANCE | CHARACTERISTIC | SYMBOL | SEE: |
|---|---|---|---|---|
| FOR INDIVIDUAL FEATURES | FORM | STRAIGHTNESS | ⏤ | 6.4.1 |
| | | FLATNESS | ⏥ | 6.4.2 |
| | | CIRCULARITY (ROUNDNESS) | ○ | 6.4.3 |
| | | CYLINDRICITY | ⌭ | 6.4.4 |
| FOR INDIVIDUAL OR RELATED FEATURES | PROFILE | PROFILE OF A LINE | ⌒ | 6.5.2 (b) |
| | | PROFILE OF A SURFACE | ⌓ | 6.5.2 (a) |
| FOR RELATED FEATURES | ORIENTATION | ANGULARITY | ∠ | 6.6.2 |
| | | PERPENDICULARITY | ⊥ | 6.6.4 |
| | | PARALLESIM | ∥ | 6.6.3 |
| | LOCATION | POSITION | ⌖ | 5.2 |
| | | CONDENTRICITY | ◎ | 5.11.3 |
| | | SYMMETRY | ⌯ | 5.13 |
| | RUNOUT | CIRCULAR RUNOUT | ↗ • | 6.7.1.21 |
| | | TOTAL RUNOUT | ⌰ • | 6.7.1.2.2 |
| • ARROWHEADS MAY BE FILLED OR NOT FILLED | | | | 3.3.1 |

**FIGURE B.8** The 14 geometric tolerances and their symbols. Reprinted from ASME Y14.3-1975 and ASME Y14.5-1994 (R2004), by permission of The American Society of Mechanical Engineers. All rights reserved.

control frame is *attached to a surface* via a leader line (e.g., the flatness tolerance associated with the screwdriver blade in the top view in Figure B.6); *placed off of an extension line from a surface* (e.g., the flatness and perpendicularity tolerances associated with the tip of the blade in the front view in Figure B.6); or *associated with the size dimension of a particular feature* (e.g., the position tolerances on the hole in the top view of Figure B.6).

The feature control frame is broken down into the three components depicted in Figure B.9. We will define the parts from left to right. The first box (1) is for the geometric characteristic symbol, which tells us what tolerance is being specified. In this case, it is position.

The second box (2) contains the actual permitted variation, or tolerance with some optional modifiers. In this particular case, the tolerance is .014 in. The diameter symbol in front of the number indicates that the tolerance zone shape is cylindrical. Figure B.10 indicates the difference between the presence and absence of a diameter symbol in defining a tolerance zone shape. A position tolerance with a diameter symbol means that the position of the item being controlled must fit inside a cylindrical tolerance zone of a

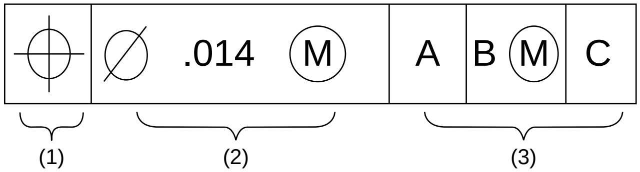



**FIGURE B.9** A feature control frame for an object that specifies the position of that object to a cylindrical tolerance zone of 0.014 in. with respect to a reference frame determined by datums A, B, and C.

diameter specified by the tolerance. Lack of a diameter symbol indicates that the position must fall between two parallel planes; the distance between those planes is defined by the tolerance specified. The tolerance material condition modifier appears after the tolerance itself and specifies the conditions under which this tolerance applies. Material condition modifiers will be described fully below.

The last set of boxes (3) contains the datum references. These define the frame of reference from which the tolerance is measured. The datum references appear in a specific order that indicates their relative importance. Note that datum references may also have material condition modifiers (described immediately below).

The feature control frame described in Figure B.9 can be read as follows, assuming it is associated with the size dimension for a hole: The hole has a permissible variation in *position* such that the center of the hole must fit within a *cylindrical tolerance zone* that is *.014 in*. in diameter when the hole is at *maximum material condition* (MMC), with respect to *datums A*, *then B*, *then C*.

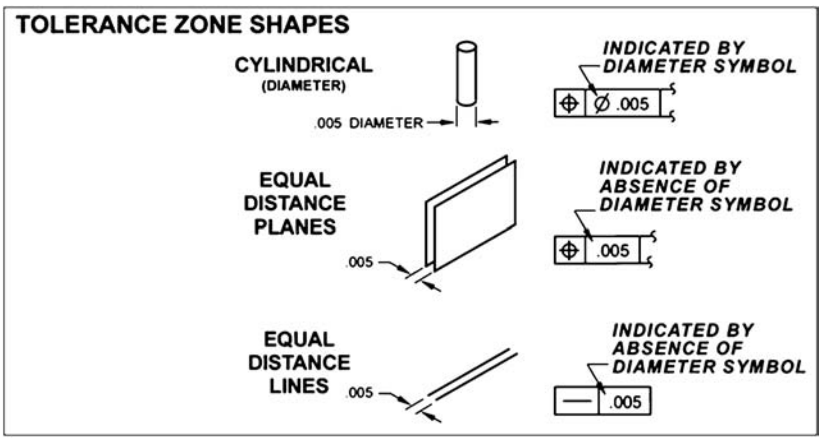



**FIGURE B.10** The tolerance zone shape depends on the type of tolerance being specified and the presence or absence of a diameter symbol before the tolerance in the feature control frame. Reprinted by permission of Technical Documentation Consultants of Arizona, Inc.

### B.2.3 Material Condition Modifiers

Three *material condition modifiers* specify the state of the feature when the tolerance is applied. These are: maximum material condition, least material condition (LMC), and regardless of feature size (RFS). It is important to know under what conditions the part is toleranced, because use of these modifiers can lead to substantial cost savings.

*Maximum material condition* is the condition in which a feature of size contains the maximum amount of material (weighs the most) within its size tolerance. MMC is indicated by the letter M enclosed in a circle, as in Figure B.9. For a hole, this means the minimum diameter specified in the size tolerance. For a cylindrical shaft, this means the maximum diameter specified by the size tolerance. For example, a hole specified as .500 ± .005 in. in diameter would have an MMC size of .495 in. in diameter.

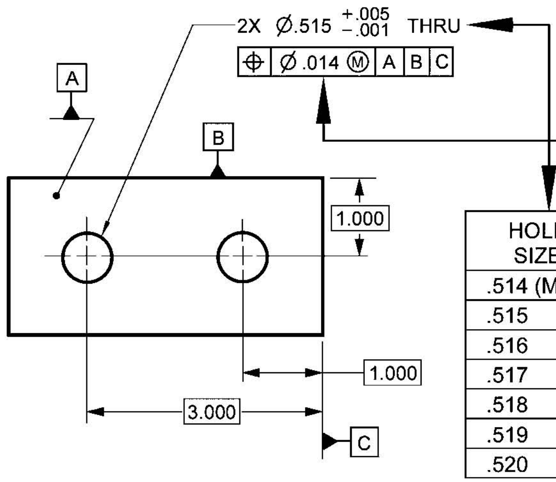
*Least material condition*, indicated by the letter L enclosed in a circle, is the condition in which a feature of size contains the least amount of material (weighs the least) within its size tolerance. The LMC size of the hole is the largest hole within the size tolerance; the LMC size of a shaft is the smallest shaft within the size tolerance. The same hole described above would have an LMC size of .505 in.

*Regardless of feature size* means just that the tolerance is applied regardless of the size of the produced part. It is indicated by the absence of either MMC or LMC modifiers.

Why would we want to use these modifiers? These modifiers are extremely useful in that they can reduce the cost of manufacturing a part substantially. They take into account the fact that if a part is produced at the extremes of its permitted size variation, there is potential for more "wiggle room" in the placement of that part. For example, if a hole is produced at its largest possible size, its position can vary more than it would if it was produced at its smallest possible size—and still match up with a mating part. The material condition modifiers enable us to have a *maximum interchangeability of parts*. This is important if we are trying to manufacture thousands of the same pieces, and we expect them all to fit together without specifying extremely tight tolerances.

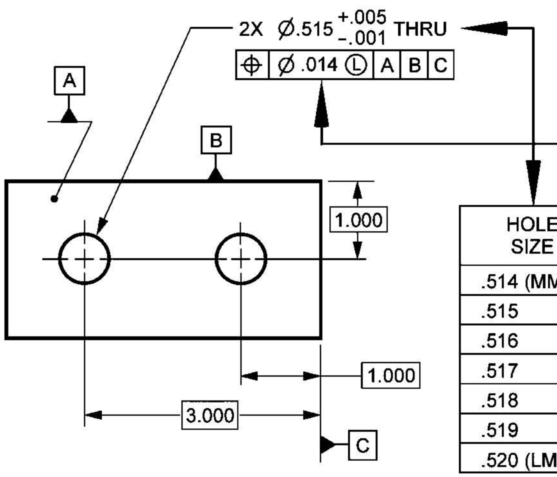
If a maximum material condition modifier is placed in a feature control frame associated with the tolerance, this means that the specified tolerance applies *only at the MMC size of the feature*. Figure B.11 shows a part with two holes controlled by a position tolerance. Since this tolerance is specified at MMC, it means that when the hole is produced at its MMC size (smallest hole, .514 in. this example), the center of the hole must fit within a cylindrical tolerance zone of diameter .014. If however, the hole is produced larger than the MMC size, additional bonus tolerance is added to the allowable position variation. The bonus tolerance added is the difference between the MMC size of the hole and the actual size of the hole. This bonus tolerance is to account for the fact that a larger hole can vary more and still line up with a mating part. Essentially, it takes into account the additive effects of the variation in size and variation in position.

We illustrate the use of a least material condition modifier in Figure B.12. The same example is shown, this time with the tolerance specified at LMC instead of MMC. In this case, when the hole is produced at its largest size, .520, the tolerance is specified as .014. As the hole gets smaller, bonus tolerance is added to allow the hole to vary more in position. The LMC modifier is used less than the MMC modifier but is useful when it is desirable that the position of a larger hole needs to be more tightly controlled, such as in the case where it is placed near the edge of a part.




| HOLE SIZE | BONUS TOLERANCE | POSITION TOLERANCE |
|---|---|---|
| .514 (MMC) | .000 | .014 |
| .515 | .001 | .015 |
| .516 | .002 | .016 |
| .517 | .003 | .017 |
| .518 | .004 | .018 |
| .519 | .005 | .019 |
| .520 | .006 | .020 |

**FIGURE B.11** Maximum material condition modifier allows additional “bonus tolerance” if part is produced at a size other than MMC. In this example, as the produced hole gets larger, the allowable variation in its position also gets larger. Reprinted by permission of Technical Documentation Consultants of Arizona, Inc.




| HOLE SIZE | BONUS TOLERANCE | POSITION TOLERANCE |
|---|---|---|
| .514 (MMC) | .006 | .020 |
| .515 | .005 | .019 |
| .516 | .004 | .018 |
| .517 | .003 | .017 |
| .518 | .002 | .016 |
| .519 | .001 | .015 |
| .520 (LMC) | .000 | .014 |

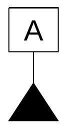
**FIGURE B.12** Least material condition modifier applies the tolerance in the case of least material condition and allows bonus tolerance for parts produced (in the case of a hole) smaller than the LMC size. This modifier is used much less frequently than the MMC modifier and can help to control hole position if the hole is located close to the edge of a part. Reprinted by permission of Technical Documentation Consultants of Arizona, Inc.

RFS does not take advantage of size variations in the feature and specifies the same tolerance for all cases. This condition is assumed if no material condition modifier is used on the drawing, so we should be careful not to omit these symbols! RFS should be used only if the requirements are very strict, since manufacturing parts to RFS is much more expensive.

One final note on material condition modifiers. These conditions may only be applied to *features of size*. A feature of size can be a cylinder, a slot, or a hole, for example. Material condition modifiers may not be used applied to surfaces, as there is no size associated with a surface. It wouldn't make sense, therefore, to have a material condition modifier associated with a tolerance in a flatness feature control frame applied to a surface.

### B.2.4 Datums

As we mentioned earlier, in the GD&T system datums form the reference frame from which to locate tolerance zones specified in the feature control frames. A few definitions are in order before we proceed. A *datum symbol* is used to define the datum on the drawing. A datum symbol looks like this:




Any letter, but for I and Q, may be used as a datum symbol. We must be careful about where we place the datum symbol because we want to be sure that the correct feature is specified as the datum. To specify a surface as a datum, the datum symbol may be placed off of an extension line or may be attached directly to the surface itself (Figure B.13). To specify a feature of size as a datum, the datum symbol may be placed in line with the dimension line for the feature, or it may be attached directly to a cylindrical feature in the view where it appears as a cylinder. Datum symbols may also be attached to the feature control frame associated with the feature of size (see Figure B.14).

A *datum feature* is what the datum symbol is applied to, the actual feature on the part. A *datum simulator* is the manufacturing and inspection tooling used to simulate the datum during production. Simulators can be a precise surface or precise tooling in which to place the part. Locations of holes or other features are then determined from the datum simulator instead of the irregular surface or edge of the part itself. The datum simulator for a surface is a surface that the part may be placed upon, and is typically made from granite due to its smooth surface free of irregularities. The simulator for a feature of size is usually a chuck or a vice that clamps around an external feature.

So how do we choose the datums for a particular part? Considerations should include the function of the part, the manufacturing processes to be employed, inspection processes that may be used, and the part's relationship to other parts. For a rectangular object, three datum references must be chosen to refer to three perpendicular planes (see Figure B.15). The *primary datum* (A) is listed first in the feature control frame and must make three points of contact on that surface. If our rectangular part is going to fit flush with another part, the largest contacting surface should be chosen as the primary datum. The primary datum creates a flat surface. The *secondary datum* (B) is usually the longest side, or a side in contact with a mating part, and requires two points of contact. This datum creates alignment and stability.

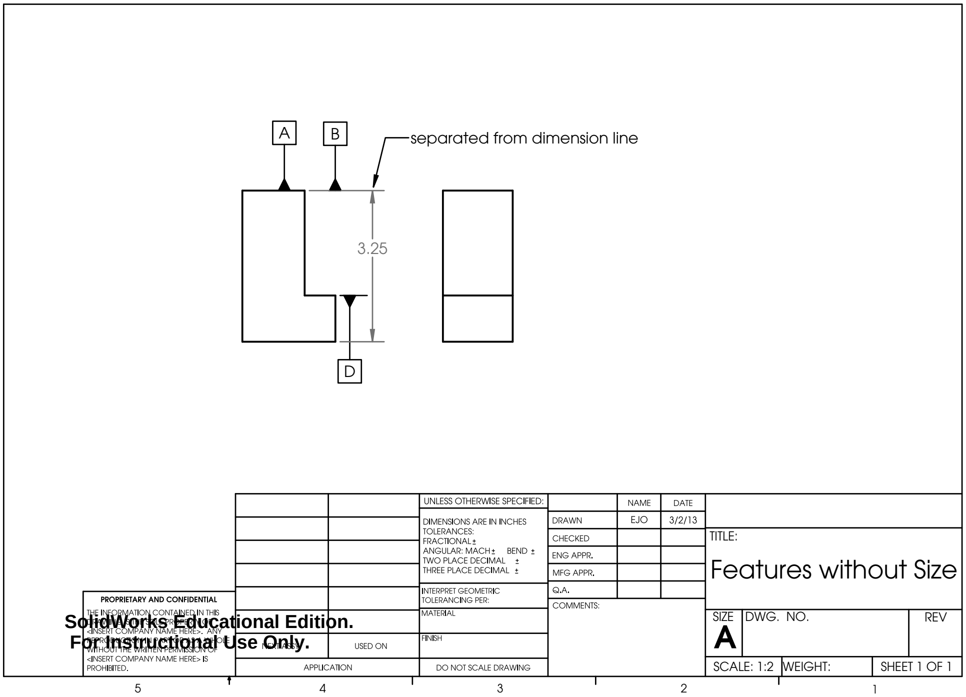



**FIGURE B.13** Specifying surfaces as datum features. The datum symbol may be placed directly on the surface or off an extension line from the surface. The datum symbol must be separated from the dimension line. Courtesy of Elizabeth J. Orwin in SolidWorks™.

The *tertiary datum* (C) is, then, the other side of the part. This datum requires one point of contact and prevents the part from sliding on datum B. In order to measure the accuracy of a part for testing, or to machine a hole located with respect to these datums, the part must first be set down on datum A, slid over to make contact with datum B, and then slid until it makes contact with datum C, while maintaining contact with datums A and B.

For a cylindrical object, two datum references are required (see Figure B.16) One reference is the surface, the other is the axis determined by a particular feature of size. In Figure B.16, the primary datum D is the bottom surface; it establishes a flat surface with three points of contact. The secondary datum E is established by the axis of the cylindrical part. This axis establishes two planes that bisect at the axis. To measure or locate from this datum, the part must be contacted by a chuck at three points. To make this particular part, the cylinder would be placed on a precise surface, and grabbed by a chuck to establish the axis, and then the holes would be located from there.

Many parts have large irregular surfaces that are not flat surfaces and not cylindrical parts. For these parts, it is impractical to define datums as we have described above. In these cases, it is permissible to identify datums using points, lines or areas instead of a whole surface. These points are called *datum targets* and specify where the workpiece contacts the

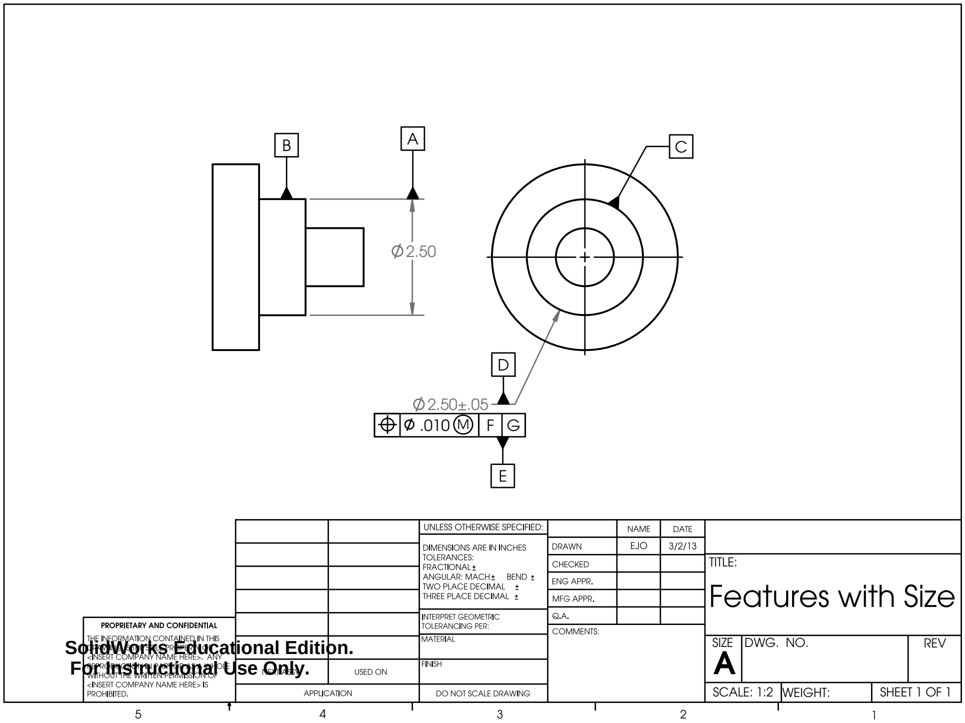



**FIGURE B.14** Specifying features of size (such as holes or shafts) as datum features. The datum symbol may be placed in line with the size dimension of the feature, placed on the feature itself, or associated with the feature control frame. All the datum symbols on this drawing indicate the .250 diameter cylindrical shaft as the datum. Courtesy of Elizabeth J. Orwin in SolidWorks™.

tooling during manufacturing and inspection. A datum target is indicated by an “X” on the drawings, and the datum reference symbols are defined in circles. Since a whole surface is not in contact, the points of contact are typically numbered “A1,” “A2,” and so on. Figure B.17 shows a hammer handle from Harvey Mudd’s introductory engineering design course with X’s marking the datum targets on this irregularly shaped handle surface. The profile of the surface is then permitted to vary with respect to those points.

One last note on datums and datum references. It is important to understand that not all types of tolerance specifications require a datum reference. (Recall that we showed all of the geometric tolerances in Figure B.8.) Note that the first set of tolerances are form tolerances and apply to *individual features*. The column on the far left in the figure distinguishes the tolerances that apply to individual features versus ones that apply to *related features*. For example, a flatness tolerance is applied to a surface: That surface is specified to be flat, but it is not with respect to any frame of reference. It would be inappropriate to specify a datum reference in this case. In contrast, a perpendicularity tolerance specifies that some feature be perpendicular to something; this something must be defined by one or more datum references.

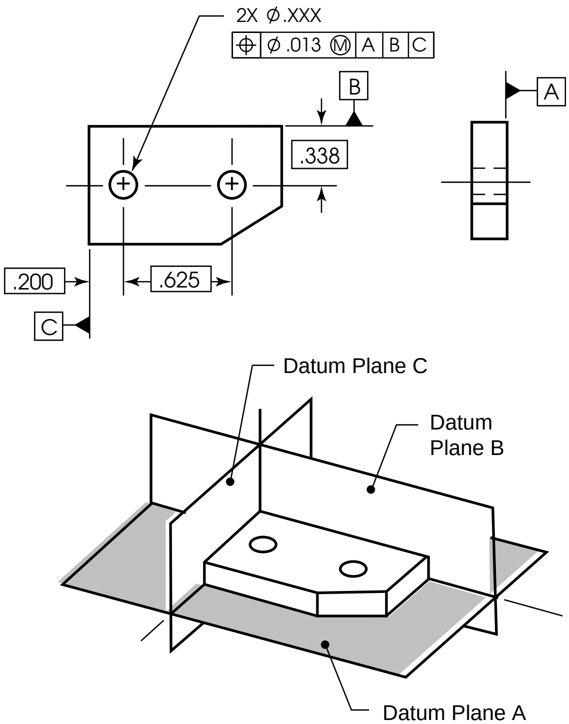



**FIGURE B.15** Specifying datums for a rectangular feature. The function of the part is important for specifying datums. The primary datum is usually chosen as the largest contacting surface. Reprinted from ASME Y14.3-1975 and ASME Y14.5-1994 (R2004), by permission of The American Society of Mechanical Engineers. All rights reserved.

## B.2.5 Position Tolerance

We have used position tolerance examples throughout the foregoing discussion, and we will now *put the pieces together* using the illustrative example shown in Figure B.18. The part depicted in this drawing has a specified position tolerance on the hole. The specified tolerance defines a cylindrical tolerance zone (note the diameter symbol) .100 in. in diameter that extends through the part. The hole axis may be tilted, but it must fit within that tolerance zone. Since MMC is specified, this tolerance is required to be met at MMC only, that is, at the smallest hole size. As the hole gets larger, bonus tolerance is added, making the cylindrical tolerance zone for the hole axis larger as the hole gets larger. The theoretical center of the hole is located at specified distances from the datums; these specified distances are called out using basic dimensions (boxed). To make this part, the stock piece would be placed on a datum simulator surface (datum A), pushed up against datum simulator surface B, and slid along to make contact with datum simulator surface C. The theoretical center of the hole would then be located from the datum simulators surfaces B and C using the basic dimensions on the drawing. Any drawing that has tolerances

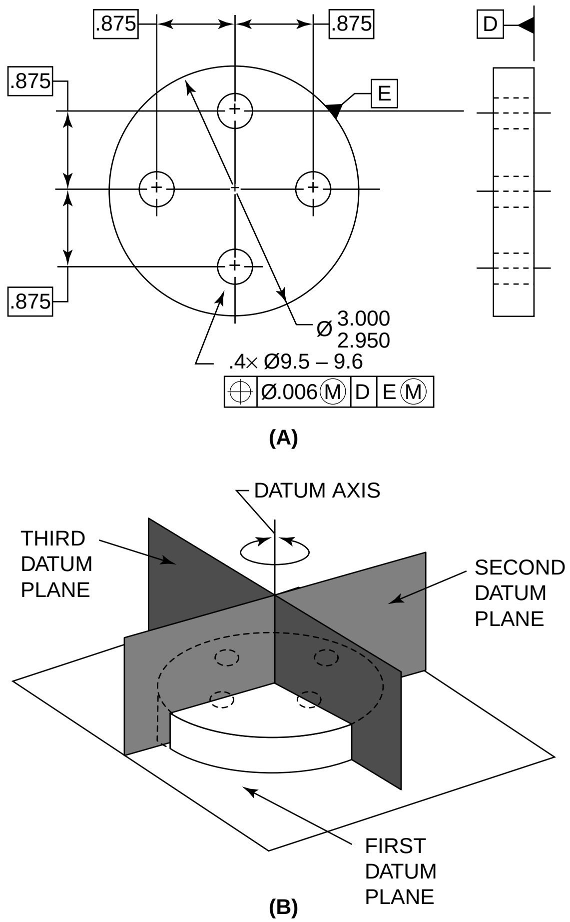



**FIGURE B.16** Specifying datums for a cylindrical feature. The primary datum is usually chosen as the flat surface to stabilize the part. The secondary datum is the axis described by the cylindrical feature. Reprinted from ASME Y14.3-1975 and ASME Y14.5-1994 (R2004), by permission of The American Society of Mechanical Engineers. 

specified by a feature control frame will have basic dimensions defining distances between the datums and the theoretical position of the tolerance zone.

## B.2.6 Fasteners

How do we know how to specify position tolerance zones such that we can fasten two parts together? How do we ensure that fasteners will always fit? There are three types of fastener conditions:

- *Floating fasteners*: The fasteners pass through holes on two or more parts and are fastened with a nut on the other side. The fasteners do not need to come into contact with the part.

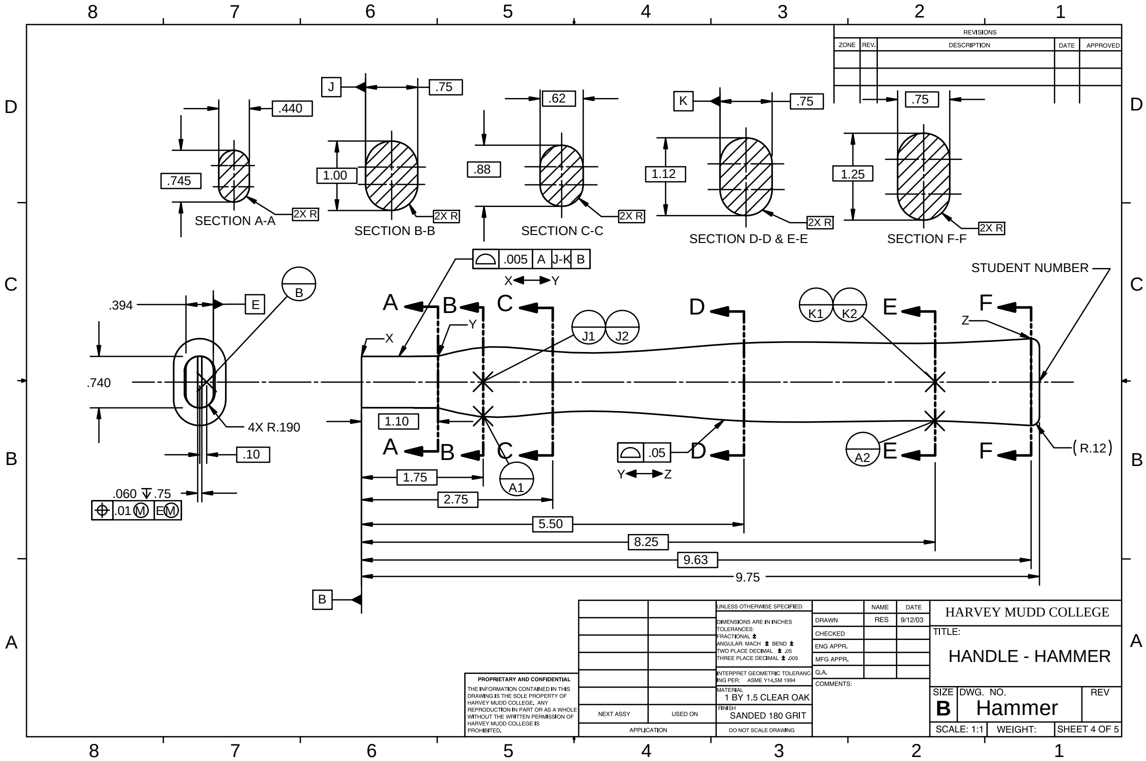



**FIGURE B.17** A drawing of the hammer handle. Note the use of datum targets (marked as "X's" on the drawing) instead of a datum feature. Courtesy of R. Erik Spjut.

- *Fixed fasteners*: One of the two (or more) parts involved is tapped or press fit (fixed) and the other has a clearance hole. The fixed part fixes the location of the fastener.
- A *double fixed fastener*: Here both holes are fixed. This gives zero position tolerance at RFS and should be avoided due to expense.

How do we calculate the position tolerance of any given hole in two parts that are to be fastened together? For a floating fastener condition, we first determine the MMC size of the hole (smallest hole, H) and the MMC size of the fastener (largest fastener, F). The difference between these two numbers gives the amount of clearance available in the worst-case scenario, when both fastener and hole are at MMC. The amount of tolerance in this case is simply this difference, that is, the tolerance $T = H - F$. For a floating fastener condition, this tolerance is applied to the position of the holes on both parts. For a fixed fastener condition, the tolerance is calculated in the same way, but now the tolerance must be distributed over the two parts. The rule of thumb is to give 60–70% of the allowable tolerance to the fixed/threaded part.

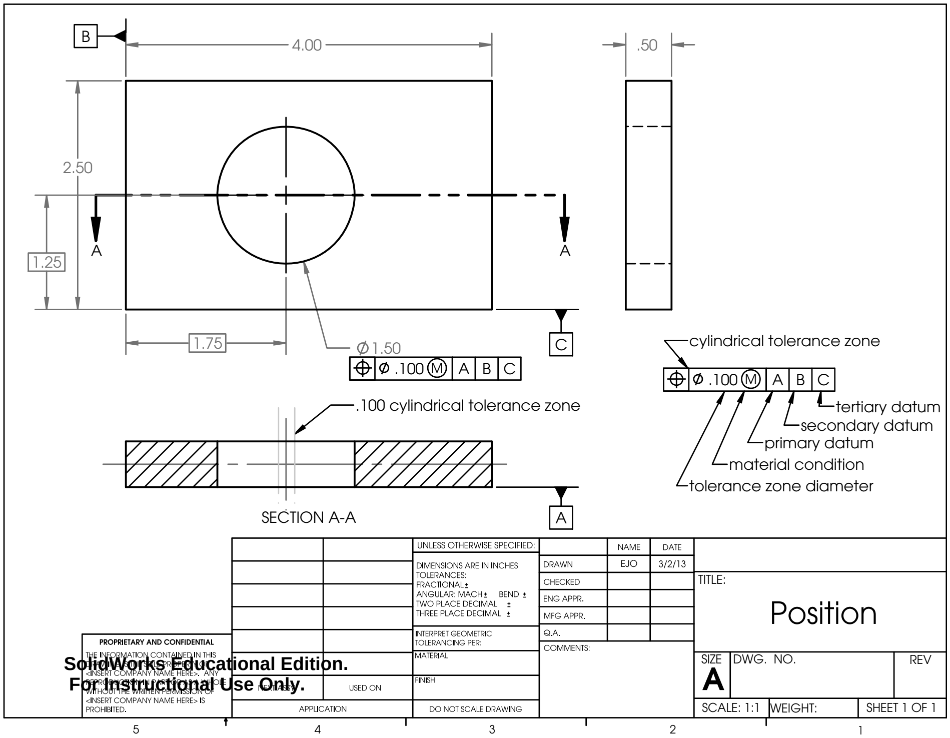



**FIGURE B.18** Putting it all together: tolerancing the true position. Courtesy of Elizabeth J. Orwin in SolidWorks™.

## B.3 HOW DO I KNOW MY PART MEETS THE SPECIFICATIONS IN MY DRAWING?

All manufactured parts need to be evaluated to make sure that they are within the specifications. A *coordinate measurement machine* (CMM) is a device that can be programmed to examine a specific part for its adherence to the geometric tolerances specified on the drawings. Often companies will invest in one of these systems if they are manufacturing a large number of similar parts and need to know if each part meets the requirements. Figure B.19 is a photograph of the hammer, described in the technical drawing in Figure B.17, mounted on the CMM system at HMC. Note that the points at which the tooling makes contact with the hammer correspond to the datum target points we saw in the drawing. This system is used to grade student-machined tools at HMC, but is more widely used in industry for quality control of manufactured parts.

There is another, much cheaper, way of evaluating parts that fit together called *functional gaging*. As the name hints, this method evaluates the *function* of a given part, that is, will a part fit with its intended mating part? This approach further illustrates the power of the material condition modifiers. In order to understand functional gaging, we

Photo by Elizabeth J. Orwin

**FIGURE B.19** The hammer described by the drawing in Figure B.17 on the coordinate measurement machine. Note the ruby tip used to measure the part and the precise granite surface of the machine. Note also that the tooling used to hold the hammer while it is being tested make contact with the part at the datum target locations specified in the drawing.

must first define another term: The *virtual condition* of a given feature is the combined effect of the size tolerance and the geometric tolerance on the part. If we imagine an external feature, such as a cylindrical shaft, the virtual condition is the MMC size of the shaft plus the geometric tolerance. It is the largest possible shaft with the largest possible variation in position, making it the worst-case scenario for the external part fitting into a mating part. For an internal feature, such as a hole, the virtual condition is also the worst-case scenario, the MMC size of the hole (smallest hole) minus the geometric tolerance. The virtual conditions of the two parts must match in order to ensure that two parts, specified on different drawings and manufactured to specifications, will always fit together. If the virtual condition of a hole in one part is matched to the virtual condition of a shaft in another part meant to fit together, this will ensure maximum interchangeability of parts. The result is powerful. It means that all parts that meet the drawing specifications will be interchangeable, that is, the two parts need not be made specifically to fit together. This clearly offers large cost savings in the manufacturing of parts.

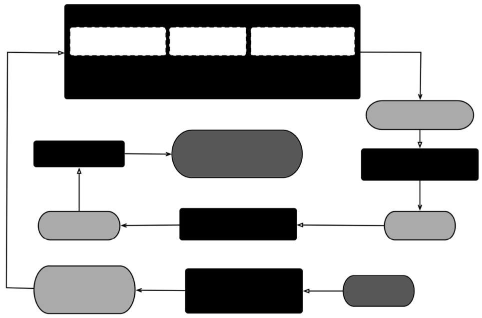
Now, back to functional gaging. Virtual condition matching allows us to use this much cheaper way to evaluate manufactured parts. In our hole/shaft example above, we could simply manufacture one part with the shaft made at virtual condition and then use that part to evaluate the potentially hundreds of matching parts. The manufactured part with the hole to be tested would simply be placed over the shaft at virtual condition: If the hole goes over the shaft, the part is good; if not, the part is thrown out.

## B.4 NOTES

*Section B.1*: Our brief overview of the basics of dimensioning and tolerancing has relied on the information found in ASME (1994), TDCA (1996), Plantenberg (2010), and Wilson (2005).

APPENDIX C

# EXERCISES



THIS APPENDIX collects in one place exercises for many of the chapters and topics covered in the main body. The authors do not use these exercises in Harvey Mudd's first-year design course. Instead, we use a faculty-generated project and a reverse engineering project as our means for teaching the design vocabulary and processes that are the meat of this book. However, for those instructors who want some additional problems for their own courses, we offer the following collection.

## C.1 EXERCISES FOR CHAPTER 1

**1.1** Define engineering design in your own words.

**1.2** List at least three questions you would ask if you were, respectively, a user (purchaser), a client (manufacturer), or a designer who was about to undertake the design of a portable electric guitar.

**1.3** List at least three questions you would ask if you were, respectively, a user (purchaser), a client (manufacturer), or a designer who was about to undertake the design of a greenhouse for a tropical climate.

**1.4** Suppose you are working for a start-up company that is designing a very new and innovative product. Does the client–user–designer model still apply? Who is your client in this case? Since you don't have any customers yet, who are the users and how can we capture their perspectives?

**1.5** Much of management may be said to be goal directed. Explain how this description is exemplified by the three Ss of management defined in Section 1.4.

## C.2 EXERCISES FOR CHAPTER 2

**2.1** When would you be likely to use a descriptive model of the design process?

**2.2** When would you be likely to use a prescriptive model of the design process?

**2.3** How does feedback play into the spiral model of the design process? Put another way, where does feedback come from, and how might the designer use it to improve the design outcome?

## C.3 EXERCISES FOR CHAPTER 3

**3.1** Explain the differences between biases, implied solutions, constraints, and objectives.

**3.2** Your design team has been given the problem statement shown below. Identify any biases and implied solutions that appear in this statement.

*Design a portable electric guitar, convenient for air travelers, that sounds, looks, and feels as much as possible like a conventional electric guitar.*

Revise the problem statement so as to eliminate these biases and implied solutions.

**3.3** Your design team has been given the problem statement shown below. Identify any biases and implied solutions that appear in this statement.

*Design a greenhouse for a women's cooperative in a village located in a Guatemalan rain forest. It would enable cultivation of medicinal preventive herbs and aid the villagers' diets. It would also be used to grow flowers that can be sold to supplement villagers' income. The greenhouse must withstand very heavy daily rains and protect the plants inside. The greenhouse must be made of indigenous materials, because the villagers are poor.*

Revise the problem statement to eliminate these biases and implied solutions.

**3.4** You are the leader of the portable guitar design team. At a team meeting, someone suggests that since you are under time pressure, you should "just skip revising the original problem statement and jump into the real design work." You point out that failing to clarify the original problem statement might lead to a bad design. Give at least two examples of how this might happen in the guitar design case.

## C.4 EXERCISES FOR CHAPTER 4

**4.1** You are the leader of the portable guitar design team. A team member wants to skip clarifying the objectives for the project, since "it's pretty obvious what constitutes a good portable guitar." Respond to their idea, explaining why clarifying objectives will help get to a better design and also what might happen if you overlook objectives.

**4.2** Develop a list of objectives and an objectives tree for the portable electric guitar. (Some team members may want to play the roles of client and users for this design project.)

**4.3** Develop a list of objectives and an objectives tree for the rainforest project. (Someone will have to play the roles of client and users for this design project.)

**4.4** Correct and revise the objectives tree developed by arm support Team A, shown in Figure 4.3.

**4.5** Correct and revise the list of objectives developed by arm support Team B, shown in Table 4.10.

**4.6** Develop a set of metrics for the portable electric guitar of Exercise 4.1.

**4.7** Develop a set of metrics for the rain forest project of Exercise 4.2.

## C.5 EXERCISES FOR CHAPTER 5

**5.1** Explain the differences between constraints and objectives.

**5.2** Under what circumstances might you convert an objective into a constraint?

**5.3** Identify constraints relevant to the portable electric guitar design problem (Exercise 3.2).

**5.4** Are there objectives in the problem statement in Exercise 3.2 that might be converted to constraints?

**5.5** Identify constraints relative to the rain forest design problem (Exercise 3.3).

**5.6** Are there objectives in the problem statement in Exercise 3.3 that might be converted to constraints?

## C.6 EXERCISES FOR CHAPTER 6

**6.1** Explain the differences between functions and objectives.

**6.2** Explain the differences between metrics and specifications.

**6.3** Using each of the methods for developing functions described in Section 6.1, develop a list of the functions of the portable electric guitar of Exercise 3.2. How effective was each of these methods in developing the specific functions?

**6.4** Using each of the methods for developing functions described in Section 6.1, develop a list of the functions of the rain forest project of Exercise 3.5. How effective was each of these methods in developing the specific functions?

**6.5** Based on the results of either Exercise 6.3 or 6.4, discuss the relationship between methods of determining functions, the nature of the functions being determined, and the nature of the device or system being designed. Is the designer's level of experience also likely to affect the outcome of functional analysis?

**6.6** Do the research necessary to determine whether there are any applicable standards (e.g., safety standards, performance standards, interface standards) for the design of the portable electric guitar of Exercise 3.2.

**6.7** Describe the interfaces between the portable electrical guitar of Exercise 3.2 and, respectively, the user and the environment. How do these interfaces constrain the design?

**6.8** Developing countries often have different safety (and other) standards than are typically found in countries such as Canada and the United States. How could this affect the design of both the portable electric guitar of Exercise 3.2 and the rain forest project of Exercise 3.5?

**6.9** What are the interface design boundaries and issues for the design and installation of a new toilet for a building?

## C.7 EXERCISES FOR CHAPTER 7

**7.1** Explain what is meant by the term "design space" and discuss how the size of the design space might affect a designer's approach to an engineering design problem.

**7.2** Using the functions developed in Exercise 6.3, develop a morphological chart for the portable electric guitar.

**7.3** Select means for realizing the design of the portable electric guitar.

**7.4** Using Web-based patent lists (see Section 7.2.2), develop a list of patents that are applicable to the portable electric guitar.

**7.5** Using the functions developed in Exercise 6.4, develop a morphological chart for the rain forest project.

**7.6** Organize and apply a process for selecting means for realizing the design of the rain forest project.

**7.7** Using Web-based patent lists (see Section 7.2.2), develop a list of patents that are applicable to the rain forest project.

**7.8** Describe an acceptable proof of concept for the rain forest project. Would a prototype be appropriate for this project? If so, what would be the nature of such a prototype?

## C.8 EXERCISES FOR CHAPTER 8

**8.1** In our discussion of how to choose among alternatives, we noted that "our information necessarily reflects a fair amount of subjectivity." Does this mean that the information is unreliable or should be rejected? How do we reconcile this subjectivity with good engineering practice?

**8.2** Some engineers like to "sum up" the results of a numerical evaluation matrix. Why is this wrong? What sorts of mistakes might it lead to?

**8.3** Apply the three methods described in Section 8.1 for choosing among alternative designs to the portable guitar design problem.

**8.4** Apply the three methods described in Section 8.1 for choosing among alternative designs to the rain forest design problem.

## C.9 EXERCISES FOR CHAPTER 9

**9.1** ***Thought Exercise*****:** Do this exercise with a partner. **Do not read the other's instructions.** *To be read by Partner #1*: You will be completing a simple paper design exercise: Design a handheld hole puncher. Design and document your item on paper in whatever manner you feel is required such that someone could interpret and manufacture it. *To be read by Partner #2*: Take the drawing from your partner. Try to answer the following questions: What is being represented? What steps are needed to manufacture the object? What materials would you use? Could you create this object in a repeatable manner from this drawing? How do parts fit together? *Both partners*: From this exercise, come up with a list of necessary components for a technical design drawing.

**9.2** Explain why having effective and accurate design drawings is particularly important in organizations that use outsourcing for manufacturing.

## C.10 EXERCISES FOR CHAPTER 10

**10.1** Describe the difference between a prototype, a model, and a proof of concept.

**10.2** At which stages in the design of a portable electric guitar would you want to use a proof of concept, a model, and a prototype?

**10.3** At which stages in the rain forest design project would you want to use a proof of concept, a model, and a prototype?

**10.4** How might your answer to Problem 10.3 change if you were developing the concept for use in many different villages or different countries? Does this tell us anything about building one-of-a-kind versus many of the something?

## C.11 EXERCISES FOR CHAPTER 11

**11.1** How can a design team determine the audience for an oral presentation?

**11.2** How does the composition of an audience affect the structure and content of a design team's presentation to that audience?

**11.3** How does a design review differ from a public presentation of project results?

**11.4** What is the difference between a rough outline and a topic sentence outline? When would you choose to use each of these outlines?

**11.5** Why is it helpful to have someone who did not write a part or section of a report review that part or section?

## C.12 EXERCISES FOR CHAPTER 12

**12.1** What is the value of the gravitational acceleration $g$ when expressed in the dimensions of furlongs and fortnights? (*Hints*: The *furlong* is a unit of length used at racetracks, and a *fortnight* is an old British term for a certain unit of time.)

**12.2** What is the constant in eq. (12.9)? Explain your answer.

**12.3** Use the basic method of dimensional analysis to determine the relationship between the natural frequency of a simple pendulum and the pendulum's mass, $m$ and the length of its weightless string of length, $l$.

**12.4** What is the effective damping coefficient of two damping elements acting in parallel?

**12.5** What is the effective capacitance of two capacitors acting in parallel?

**12.6** What is the effective inductance of two inductors acting in parallel?

**12.7** Show that the pendulum equations (12.14) and (12.16)–(12.19) are dimensionally homogeneous.

**12.8** What are the physical dimensions of the parameter $g/l$ that appears in the pendulum equations of motion, that is, eqs. (12.14) and (12.15).

**12.9** Can the results of Problem 9.8 be used to make the time, $t$ dimensionless?

**12.10** How can we use the results of Problem 9.9 and a small-angle assumption to show that eq. (12.17) does reduce to $T \cong mg$ for small angles.

**12.11** Express the battery capacities of AA and LR44 batteries in coulombs (i.e., (As)) and compare them with the (Ah) values given in Section 12.3.2.

**12.12** How would you model the *two*-subsystem model of the battery–motor–cart problem depicted in Figure 12.4(b), that is: What inputs and outputs would you include? What principles and which governing equations need to be invoked?

**12.13** How would you model the *three*-subsystem model of the battery–motor–cart problem depicted in Figure 12.4(c), that is: What inputs and outputs would you include? What principles and which governing equations need to be invoked?

**12.14** Solve eq. (12.58) for $\omega = \omega(T)$ and then plot $\omega$ versus $T$ and compare it with the characteristic curve in Figure 12.7.

**12.15** Recast eq. (12.59) for $P_{out}(T) = T\omega(T)$ and then plot $P_{out}(T)$ versus $T$ and compare it with the power curve in Figure 12.7.

**12.16** Do the reasons for selecting among wood, fiberglass, and CFRP for structural members change as the sorts of loads applied and the geometric constraints change?

**12.17** Why is fiberglass used so extensively to make commercial ladders?

## C.13 EXERCISES FOR CHAPTER 13

**13.1** Verify the present value of the lifetime cost of \$1595 for the second system designed in Section 13.2.

**13.2** Your design team has produced two alternative designs for city buses. Alternative *A* has an initial cost of \$100,000, estimated annual operating costs of \$10,000, and will require a \$50,000 overhaul after five years. Alternative *B* has an initial cost of \$150,000, estimated annual operating costs of \$5000, and will not require an overhaul after five years. Both alternatives will last ten years. If all other vehicle performance characteristics are the same, determine which bus is preferable using a discount rate of 10%.

**13.3** Would the decision reached in Exercise 13.2 change if the discount rate was 20%? What would happen if, instead, the discount rate was 15%? How do the resulting cost figures influence your assessments of the given cost estimates?

**13.4** Your design team has produced two alternative designs for greenhouses in a developing country. Alternative *A* has an initial cost of $200 and will last two years. Alternative *B* has an initial cost of $1000 and will last ten years. All other things being equal, determine which greenhouse is more economical with a discount rate of 13%. What other factors might influence this decision?

**13.5** Given the definition of a discount rate, are rates of 10–20% reasonable for projects that are intended to assist the poor in developing countries? Explain why or why not.

## C.14 EXERCISES FOR CHAPTER 14

**14.1** If you were asked to design a product for recyclability, how would you determine what that meant? In addition, what sorts of questions should you be prepared to ask and answer?

**14.2** How might DFA considerations differ for products made in large volume (e.g., the portable electric guitar) and those made in very small quantities (e.g., a greenhouse)?

**14.3** What is the reliability of the system portrayed in Figure 14.2?

**14.4** What is the reliability of the system portrayed in Figure 14.3? How does this result compare with that of Exercise 14.3? Why?

**14.5** On what basis would you choose between a single system, all of whose parts are redundant, and two copies of a system that has no redundancy?

**14.6** What are the factors to consider in an environmental analysis of the beverage container design problem? How might an environmental life cycle assessment help in addressing some of these questions?

**14.7** What are some environmental considerations that might affect the design of the greenhouse for use in a developing country?

## C.15 EXERCISES FOR CHAPTER 15

**15.1** When might a student design team find it particularly helpful to discuss the stages of group formation?

**15.2** In your role as the team leader of an engineering design team, you encounter the following situation while preparing the team's final report. Whenever Ken submits written materials documenting his work, David criticizes the writing so severely and personally that the other members of the team become quite uncomfortable.

You recognize that Ken's work product does not meet your standards, but you also consider David's approach to be counterproductive. Can you resolve this conflict constructively? How?

**15.3** Discuss the differences and similarities between leadership and membership in a design setting.

**15.4** Your engineering manager is considering using a personality test to help her in assigning teams. She asks for your thoughts about how this might be useful in the team formation process. Briefly summarize your response to her.

## C.16 EXERCISES FOR CHAPTER 16

**16.1** Explain the differences between managing design projects and building the outcomes of design projects. For example, consider the differences between designing a highway interchange and building that interchange.

**16.2** Develop a work breakdown structure (WBS) for an on-campus benefit to raise money for the homeless.

**16.3** Develop a team charter and a work breakdown structure (WBS) for a project to design a robot that will be entered into a national collegiate competition.

**16.4** Develop a schedule and a budget for the on-campus benefit of Exercise 16.2.

**16.5** Your design team has completed its work on a design project. You have been asked by the team leader to organize a postaudit review of your team's work. Explain your strategy for conducting such an audit.

## C.17 EXERCISES FOR CHAPTER 17

**17.1** Describe in your own words the difference between ethics and morals.

**17.2** Identify the stakeholders that a design team must recognize as it develops its design for the portable electric guitar. Are there obligations to these stakeholders that you should consider, and could they conflict with what your client has asked you to do?

**17.3** As an engineer testing designs for electronic components, you discover that they fail in a particular location. Subsequent investigation shows that the failures are due to a nearby high-powered radar facility. While you can shield your own designs so that they will work in this environment, you also notice that there is an adjacent nursery school. What actions, if any, should you take?

**17.4** You are considering a safety test for a newly designed device. Your supervisor instructs you not to perform this test because the relevant government regulations are silent on this aspect of the design. What actions, if any, should you take?

**17.5** As a result of previous experiences as a designer of electronic packaging, you understand a sophisticated heat-treatment process that has not been patented, although it is considered company-confidential. In a new job, you are designing beverage containers for BJIC and you believe that this heat-treatment process could be effectively used. Can you use your prior knowledge?

**17.6** With reference to Exercise 17.5, suppose your employer is a nonprofit organization that is committed to supplying food to disaster victims. Would that change the actions you might take?

**17.7** You are asked to provide a reference for a member of your design team, Jim, in connection with a job application he has filed. You have not been happy with Jim's performance, but you believe that he might do better in a different setting. While you are hopeful that you can replace Jim, you also feel obligated to provide an honest appraisal of Jim's potential. What should you do?

**17.8** With reference to Exercise 17.7, would your answer change if you knew that you could not replace Jim?

## C.18 EXERCISES FOR APPENDIX A

**AA.1** Create sketches or CAD drawings of a hollow 6 in. × 6 in. × 6 in. cube with a frame of 1/4 in. square 6061 aluminum stock and side panels of 1/16 in. thick polystyrene sheets.

**AA.2** Develop a bill of materials and a budget for the hollow cube of Exercise 9.3. (*Hint*: Visit a supply website such as *McMaster-Carr* <http://www.mcmaster.com/>.)

**AA.3** Develop a process router for the hollow cube of Exercise 9.3.

**AA.4** Create sketches or CAD drawings of a hollow cube (3 ft on a side) with a frame of 1 in. × 2 in. (nominal) furring strips and side panels of 1/4 in. thick BC plywood.

**AA.5** Develop a bill of materials and a budget for the hollow cube of Exercise 9.3. (*Hint*: Try websites such as *Lowe's* <http://www.lowes.com/>, or *Home Depot* <http://www.homedepot.com/>.)

## C.19 EXERCISES FOR APPENDIX B

**AB.1** Draw the correct third-angle projection orthographic views for a block letter "F."

**AB.2** If a dimension on a drawing has no associated tolerance, where should the machinist look to determine the permissible variation?

**AB.3** Is a datum reference needed when specifying a flatness tolerance? Why (or why not)?

**AB.4** Sketch a functional gage for the part shown in Figure 8.16 to test the position of the holes.

**AB.5** Make a technical drawing of a rectangular part 3 in. long, 2 in. wide, with a thickness of 0.5 in. The part has one hole, located 0.75 in. from the part's left side of the part and 0.75 in. from its bottom. The hole has a diameter of 0.25 in. All size dimensions can vary ±0.01 in. and the location of the hole can vary 0.005 in. Design for maximum interchangeability of parts.

# REFERENCES AND BIBLIOGRAPHY

This combined bibliography contains both references consulted in the previous editions of this book and a sampling of books on issues such as design theory, design in different disciplines, product development, project management techniques, optimization theory, applications of artificial intelligence, engineering ethics and the practice of engineering, and more. This is *not* a complete list of works: The literatures on design and project management alone are both vast and rapidly expanding. Thus, keep in mind that this list represents only the tip of a very large iceberg of published work in design and in project management. Some of the works cited are just intellectually interesting, and some are books that students in particular will find useful for project work.


N. Abram, *Measure Twice, Cut Once: Lessons from a Master Carpenter*, Little, Brown, Boston, MA, 1996.

J. L. Adams, *Conceptual Blockbusting: A Guide to Better Ideas*, Stanford Alumni Association, Stanford, CA, 1979.

H. Ahmad, B. Liu, C. Kneen, and M. Pham, *Artwork & The Tracka: Enhancing the Functional Independence of Children with Cerebral Palsy*, E4 Project Report, Department of Engineering, Harvey Mudd College, Claremont, CA, 2007.

K. Akiyama, *Function Analysis: Systematic Improvement of Quality and Performance*, Productivity Press, Cambridge, MA, 1991.

C. Alexander, *Notes on the Synthesis of Form*, Harvard University Press, Cambridge, MA, 1964.

Anon., *American National Standard for Ladders—Wood Safety Requirements*, ANSI A14.1-2000, American National Standards Institute (ANSI), Chicago, IL, 2000.

Anon., *Dimensions and Tolerancing*, ANSI Y 14.5M-1994, American National Standards Institute (ANSI), Chicago, IL, 1994.

Anon., *Goals and Priorities for Research in Engineering Design*, American Society of Mechanical Engineers, New York, NY, 1986.

Anon., *Improving Engineering Design: Designing for Competitive Advantage*, National Research Council, National Academy Press, Washington, D.C., 1991.

Anon., *Managing Projects and Programs*, The Harvard Business Review Book Series, Harvard Business School Press, Cambridge, MA, 1989.

Anon., *Technical Documentation Consultants of America*, Ridgecrest, California, 1996.

E. K. Antonsson and J. Cagan, *Formal Engineering Design Synthesis*, Cambridge University Press, New York, 2001.

H. Arendt, *Eichmann in Jerusalem: A Report on the Banality of Evil*, Viking Press, New York, NY, 1963.

J. S. Arora, *Introduction to Optimum Design*, McGraw-Hill, New York, NY, 1989.

K. J. Arrow, *Social Choice and Individual Values*, 1st ed., John Wiley & Sons, Inc., New York, 1951.

M. F. Ashby, *Materials Selection in Mechanical Design*, 2nd ed., Butterworth Heinemann, Oxford, England, 1999.

W. Asimow, *Introduction to Design*, Prentice-Hall, Englewood Cliffs, NJ, 1962.

R. Attarian, N. Hasegawa, J. Osgood, and A. Lee, *Design and Implementation of an Arm Support Device for a Danbury School Student with Cerebral Palsy*, E4 Project Report, Department of Engineering, Harvey Mudd College, Claremont, CA, 2007.

A. B. Badiru, *Project Management in Manufacturing and High Technology Operations*, John Wiley & Sons, Inc., New York, 1996.

K. M. Bartol and D. C. Martin, *Management*, 2nd ed., McGraw-Hill, New York, NY, 1994.

Barton-Aschman Associates, *North Area Terminal Study*, Technical Report, Barton-Aschman Associates, Evanston, IL, August 1962.

Louis Berger, *Central Artery North Area Project*, Interim Report, Louis Berger & Associates, Cambridge, MA, 1981.

R. Best, M. Honda, J. Karras, and A. Kurtis, *Design of Arm Restraint for Student with Cerebral Palsy*, E4 Project Report, Department of Engineering, Harvey Mudd College, Claremont, CA, 2007.

G. Boothroyd and P. Dewhurst, *Product Design for Assembly*, Boothroyd Dewhurst Inc., Wakefield, RI, 1989.

T. Both, G. Breed, C. Stratton, and K. V. Horn, *Micro Laryngeal Surgery: An Instrument Stabilizer*, E4 Project Report, Department of Engineering, Harvey Mudd College, Claremont, CA, 2000.

C. L. Bovee, M. J. Houston, and J. V. Thill, *Marketing*, 2nd ed., McGraw-Hill, New York, NY, 1995.

C. L. Bovee, J. V. Thill, M. B. Word, and G. P. Dovel, *Management*, McGraw-Hill, New York, NY, 1993.

E. T. Boyer, F. D. Meyers, F. M. Croft Jr., M. J. Miller, and J. T. Demel, *Technical Graphics*, John Wiley & Sons, Inc., New York, NY, 1991.

D. C. Brown, "Design," in S. C. Shapiro (Editor), *Encyclopedia of Artificial Intelligence*, 2nd ed., John Wiley & Sons, Inc., New York, NY, 1992.

D. C. Brown and B. Chandrasekaran, *Design Problem Solving*, Pitman/Morgan Kaufmann, London/Los Altos, CA, 1989.

L. L. Bucciarelli, *Designing Engineers*, MIT Press, Cambridge, MA, 1994.

S. Carlson Skalak, H. Kemser, and N. Ter-Minassian, "Defining a Product Development Methodology with Concurrent Engineering for Small Manufacturing Companies," *Journal of Engineering Design*, *8* (4), 305–328, December 1997.

A. D. S. Carter, *Mechanical Reliability*, Macmillan, London, England, 1986.

S. Chan, R. Ellis, M. Hanada, and J. Hsu, *Stabilization of Microlaryngeal Surgical Instruments*, E4 Project Report, Department of Engineering, Harvey Mudd College, Claremont, CA, 2000.

J. Corbett, M. Dooner, J. Meleka, and C. Pym, *Design for Manufacture: Strategies, Principles and Techniques*, Addison-Wesley, Wokingham, England, 1991.

R. D. Coyne, M. A. Rosenman, A. D. Radford, M. Balachandran, and J. S. Gero, *Knowledge-Based Design Systems*, Addison-Wesley, Reading, MA, 1990.

E. F. Crawley, J. Malmqvist, S. Ostlund, and D. R. Brodeur, *Rethinking Engineering Education: The CDIO*, Springer Science + Business Media, New York, NY, 2007.

N. Cross, *Engineering Design Methods*, 2nd ed., John Wiley & Sons, Ltd, Chichester, England, 1994.

M. Davis, "Reflections on the History of Engineering in the United States: A Preface to Engineering Ethics," Paper given at GTE Lecture at the Center for Academic Ethics & College of Engineering at Wayne State University, November 19, 1992; online at http://ethics.iit.edu/resources/eethics.html, accessed April 15, 2007.

O. L. de Weck, D. Roos, and C. L. Magee, *Engineering Systems: Meeting Human Needs in a Complex Technological World*, MIT Press, Cambridge, MA, 2011.

M. L. Dertouzos, R. K. Lester, R. M. Solow, and the MIT Commission on Industrial Productivity, *The Making of America: Regaining the Productive Edge*, MIT Press, Cambridge, MA, 1989.

G. E. Dieter and L. C. Schmidt, *Engineering Design*, McGraw-Hill, New York, 2012.

J. R. Dixon, *Design Engineering: Inventiveness, Analysis, and Decision Making*, McGraw-Hill, New York, NY, 1966.

J. R. Dixon, "Engineering Design Science: The State of Education," *Mechanical Engineering*, *113* (2), February 1991a.

J. R. Dixon, "Engineering Design Science: New Goals for Education," *Mechanical Engineering*, *113* (3), March 1991b.

J. R. Dixon and C. Poli, *Engineering Design and Design for Manufacturing*, Field Stone Publishers, Conway, MA, 1995.

C. L. Dym (Editor), *Applications of Knowledge-Based Systems to Engineering Analysis and Design*, American Society of Mechanical Engineers, New York, NY, 1985.

C. L. Dym, Letter to the Editor, *Mechanical Engineering*, *114* (8), August 1992.

C. L. Dym, "The Role of Symbolic Representation in Engineering Education," *IEEE Transactions on Education*, *35* (2), March 1993a.

C. L. Dym, *E4 (Engineering Projects) Handbook*, Department of Engineering, Harvey Mudd College, Claremont, CA, Spring 1993b.

C. L. Dym, *Engineering Design: A Synthesis of Views*, Cambridge University Press, New York, NY, 1994a.

C. L. Dym, "Teaching Design to Freshmen: Style and Content," *Journal of Engineering Education*, *83* (4), 303–310, October 1994b.

C. L. Dym, *Structural Modeling and Analysis*, Cambridge University Press, New York, 1997a.

C. L. Dym (Editor), *Computing Futures in Engineering Design*, Harvey Mudd College, Claremont, CA, 1997b.

C. L. Dym (Editor), *Designing Design Education for the 21st Century*, Harvey Mudd College, Claremont, CA, 1999.

C. L. Dym, *Principles of Mathematical Modeling*, 2nd ed., Elsevier Academic Press, New York, NY, 2004.

C. L. Dym, "Basic Elements of Mathematical Modeling," in P. Fishwick (Editor), *CRC Handbook of Dynamic System Modeling*, CRC Press, Boca Raton, FL, pp. 5.1–5.20, 2007.

C. L. Dym, A. M. Agogino, D. D. Frey, O. Eris, and L. J. Leifer, "Engineering Design Thinking, Teaching and Learning," *Journal of Engineering Education*, *94* (1), 103–120, January 2005.

C. L. Dym and D. C. Brown, *Engineering Design: Representation and Reasoning*, 2nd ed., Cambridge University Press, New York, NY, 2012.

C. L. Dym and R. E. Levitt, *Knowledge-Based Systems in Engineering*, McGraw-Hill, New York, NY, 1991.

C. L. Dym and L. Winner (Editor), *Social Dimensions of Engineering Design*, Harvey Mudd College, Claremont, CA, 2001.

C. L. Dym, W. H. Wood, and M. J. Scott, "Rank Ordering Engineering Designs: Pairwise Comparison Charts and Borda Counts," *Research in Engineering Design*, *13* (4), 236–242, 2003.

C. E. Ebeling, *An Introduction to Reliability and Maintainability Engineering*, Waveland Press, Long Grove, IL, 2010.

D. L. Edel Jr., (Editor), *Introduction to Creative Design*, Prentice-Hall, Englewood Cliffs, NJ, 1967.

K. S. Edwards Jr., and R. B. McKee, *Fundamentals of Mechanical Component Design*, McGraw-Hill, New York, NY, 1991.

K. A. Ericsson and H. A. Simon, *Protocol Analysis: Verbal Reports as Data*, MIT Press, Cambridge, MA, 1984.

A. Ertas and J. C. Jones, *The Engineering Design Process*, 2nd ed., John Wiley & Sons, Inc., New York, NY, 1996.

D. L. Evans (Coordinator), "Special Issue: Integrating Design throughout the Curriculum," *Engineering Education*, *80* (5), 1990.

J. H. Faupel, *Engineering Design*, John Wiley & Sons, Inc., New York, NY, 1964.

L. Feagan, T. Galvani, S. Kelley, and M. Ong, *Device for Microlaryngeal Instrument Stabilization*, E4 Project Report, Department of Engineering, Harvey Mudd College, Claremont, CA, 2000.

A. Feenberg, *Questioning Technology*, Routledge, New York, 1999.

J. Fortune and G. Peters, *Learning From Failure—The Systems Approach*, John Wiley & Sons, Ltd, Chichester, UK, 1995.

R. L. Fox, *Optimization Methods for Engineering Design*, Addison-Wesley, Reading, MA, 1971.

M. E. French, *Conceptual Design for Engineers*, 2nd ed., Design Council Books, London, England, 1985.

M. E. French, *Form, Structure and Mechanism*, MacMillan, London, England, 1992.

D. C. Gause and G. M. Weinberg, *Exploring Requirements: Quality before Design*, Dorset House Publishing, New York, NY, 1989.

J. S. Gero (Editor), *Design Optimization*, Academic Press, Orlando, FL, 1985.

J. S. Gero (Editor), *Proceedings of AI in Design '92*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1992.

J. S. Gero (Editor), *Proceedings of AI in Design '94*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1994.

J. S. Gero (Editor), *Proceedings of AI in Design '96*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996.

M. P. Glazer and P. M. Glazer, *The Whistleblowers: Exposing Corruption in Government and Industry*, Basic Books, New York, NY, 1989.

G. L. Glegg, *The Design of Design*, Cambridge University Press, Cambridge, England, 1969.

G. L. Glegg, *The Science of Design*, Cambridge University Press, Cambridge, England, 1973.

G. L. Glegg, *The Selection of Design*, Cambridge University Press, Cambridge, England, 1972.

T. J. Glover, *Pocket Ref*, Sequoia Publishing, Littleton, CO, 1993.

B. A. Goetsch, J. A. Nelson, and W. S. Chalk, *Technical Drawing*, 4th ed., Delmar Publishers, Albany, New York, 2000.

S. H. Goldstein and R. A. Rubin, "Engineering Ethics," *Civil Engineering*, *66* (10), 41–43, October 1996.

P. Graham (Editor), *Mary Parker Follett—Prophet of Management: A Celebration of Writings From the 1920s*, Harvard Business School Press, Boston, MA, 1996.

P. Gutierrez, J. Kimball, B. Maul, A. Thurston, and J. Walker, *Design of a Chicken Coop*, E4 Project Report, Department of Engineering, Harvey Mudd College, Claremont, CA, 1997.

C. Hales, *Managing Engineering Design*, Longman Scientific & Technical, Harlow, England, 1993.

J. Harr, *A Civil Action*, Vintage Books, New York, NY, 1995.

J. R. Hauser and D. Clausing, "The House of Quality," *Harvard Business Review*, 63–73, May–June 1988.

S. I. Hayakawa, *Language in Thought and Action*, 4th ed., Harcourt Brace Jovanovich, San Diego, CA, 1978.

R. T. Hays, "Value Management," in W. K. Hodson (Editor), *Maynard's Industrial Engineering Handbook*, 4th ed., McGraw-Hill, New York, NY, 1992.

G. H. Hazelrigg, *Systems Engineering: An Approach to Information-Based Design*, Prentice Hall, Upper Saddle River, NJ, 1996.

G. H. Hazelrigg, "Validation of Engineering Design Alternative Selection Methods," unpublished manuscript, courtesy of the author, 2001

J. Heskett, *Industrial Design*, Thames and Hudson, London, 1980.

R. S. House, *The Human Side of Project Management*, Addison-Wesley, Reading, MA, 1988.

V. Hubka, M. M. Andreasen, and W. E. Eder, *Practical Studies in Systematic Design*, Butterworths, London, England, 1988.

V. Hubka and W. E. Eder, *Design Science*, Springer-Verlag, London, England, 1996.

B. Hyman, *Topics in Engineering Design*, Prentice Hall, Englewood Cliffs, NJ, 1998.

D. Jain, G. P. Luth, H. Krawinkler and K. H. Law, *A Formal Approach to Automating Conceptual Structural Design*, Technical Report No. 31, Center for Integrated Facility Engineering, Stanford University, Stanford, CA, 1990.

F. D. Jones, *Ingenious Mechanisms*: Vols. *1–3*, The Industrial Press, New York, NY, 1930.

J. C. Jones, *Design Methods*, Wiley-Interscience, Chichester, UK, 1992.

J. Juran, *Quality Control Handbook*, 3rd ed., McGraw-Hill, New York, 1979.

D. Kaminski, "A Method to Avoid the Madness," *The New York Times*, 3 November 1996.

H. Kerzner, *Project Management: A Systems Approach to Planning, Scheduling and Controlling*, Van Nostrand Reinhold, New York, NY, 1992.

D. S. Kezsbom, D. L. Schilling, and K. A. Edward, *Dynamic Project Management: A Practical Guide for Managers & Scientists*, John Wiley & Sons, Inc., New York, NY, 1989.

A. Kusiak, *Engineering Design: Products, Processes and Systems*, Academic Press, San Diego, CA, 1999.

M. Levy and M. Salvadori, *Why Buildings Fall Down*, Norton, New York, NY, 1992.

E. E. Lewis, *Introduction to Reliability Engineering*, John Wiley & Sons, Inc., New York, NY, 1987.

P. Little, *Improving Railroad Car Reliability Using a New Opportunistic Maintenance Heuristic and Other Information System Improvements*, Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA, 1991.

P. Little, D. Barney, and R. Hink, "Living Up to the Code: Engineering as Political Judgment," *International Journal of Engineering Education*, *24* (2), 314–327, March 2008.

P. Little, R. Hink, and D. Barney, "Living Up to the Code: Engineering as Political Judgment," *International Journal of Engineering Education*, *24* (2), 314–327, 2007.

M. W. Martin and R. Schinzinger, *Ethics in Engineering*, 3rd ed., McGraw-Hill, New York, NY, 1996.

Massachusetts Department of Public Works, *North Terminal*, Draft Environmental Impact Report (Section 4(F) and Section 106 Statements), Massachusetts Department of Public Works, Boston, MA, 1974.

J. McGourty and K. P. DeMeuse, *The Team Developer: An Assessment and Skill Building Program*, Instructor's Manual, John Wiley & Sons, Inc. New York, 2001.

B. McKibben, *Deep Economy: The Wealth of Communities and the Durable Future*, Times Press, New York, 2007.

R. L. Meehan, *Getting Sued and Other Tales of the Engineering Life*, The MIT Press, Cambridge, MA, 1981.

J. R. Meredith and S. J. Mantel Jr., *Project Management: A Managerial Approach*, John Wiley & Sons, Inc., New York, NY, 1995.

F. C. Mish (Editor), *Webster's Tenth New Collegiate Dictionary*, Merriam-Webster, Springfield, MA, 1993.

J. Morgenstern, "The Fifty-Nine-Story Crisis," *The New Yorker*, 29 May 1995.

T. T. Nagle, *The Strategy and Tactics of Pricing*, Prentice-Hall, Englewood Cliffs, NJ, 1987.

A. Newell and H. A. Simon, *Human Problem Solving*, Prentice-Hall, Englewood Cliffs, NJ, 1972.

R. L. Norton, *Machine Design: An Integrated Approach*, 3rd ed., Prentice Hall, Upper Saddle River, NJ, 2004.

G. D. Oberlander, *Project Management for Engineering and Construction*, McGraw-Hill, New York, NY, 2000.

K. N. Otto, "Measurement Methods for Product Evaluation," *Research in Engineering Design*, *7*, 86–101, 1995.

K. N. Otto and K. L. Wood, *Product Design: Techniques in Reverse Engineering and New Product Development*, Prentice Hall, Upper Saddle River, NJ, 2001.

G. Pahl and W. Beitz, *Engineering Design: A Systematic Approach*, 2nd ed., Springer, London, England, 1996.

A. Palladio, *The Four Books of Architecture*, Dover, New York, NY, 1965.

Y. C. Pao, *Elements of Computer-Aided Design and Manufacturing*, John Wiley & Sons, Inc., New York, NY, 1984.

P. Y. Papalambros and D. J. Wilde, *Principles of Optimal Design: Modeling and Computation*, 2nd ed., Cambridge University Press, Cambridge, England, 2000.

T. E. Pearsall, *The Elements of Technical Writing*, Allyn & Bacon, Needham Heights, MA, 2001.

H. Petroski, *Design Paradigms*, Cambridge University Press, New York, NY, 1994.

H. Petroski, *Engineers of Dreams*, Alfred A. Knopf, New York, NY, 1995.

H. Petroski, *To Engineer Is Human*, St. Martin's Press, New York, NY, 1985.

W. S. Pfeiffer, *Pocket Guide to Technical Writing*, Prentice Hall, Upper Saddle River, NJ, 2001.

L. Phlips, *The Economics of Price Discrimination*, Cambridge University Press, Cambridge, England, 1985.

K. Plantenberg, *Engineering Graphics Essentials*, 4th ed., SDC Publications, Mission, KS, 2010.

S. Pugh, *Total Design: Integrated Methods for Successful Product Engineering*, Addison-Wesley, Wokingham, England, 1991.

H. E. Riggs, *Financial and Cost Analysis for Engineering and Technology Management*, John Wiley & Sons, Inc., New York, 1994.

J. L. Riggs and T. M. West, *Essentials of Engineering Economics*, McGraw-Hill, New York, NY, 1986.

E. S. Rubin, *Introduction to Engineering and the Environment*, McGraw-Hill, New York, 2001.

M. D. Rychener (Editor), *Expert Systems for Engineering Design*, Academic Press, Boston, MA, 1988.

D. G. Saari, *Basic Geometry of Voting*, Springer-Verlag, New York, 1995.

D. G. Saari, "Bad Decisions: Experimental Error or Faulty Decision Procedures," unpublished manuscript, courtesy of the author, 2001a.

D. G. Saari, *Decisions and Elections: Explaining the Unexpected*, Cambridge University Press, New York, 2001b.

M. Salvadori, *Why Buildings Stand Up*, McGraw-Hill, New York, NY, 1980.

A. Samuel, *Make and Test Projects in Engineering Design*, Springer-Verlag, London, UK, 2006.

Y. Saravanos, J. Schauer, and C. Wassman, *Sliding Fulcrum Stabilizer*, E4 Project Report, Department of Engineering, Harvey Mudd College, Claremont, CA, 2000.

D. A. Schon, *The Reflective Practitioner*, Basic Books, New York, NY, 1983.

D. Schroeder, "Little Land Bruisers," *Car and Driver*, 96–109, May 1998.

R. G. Schroeder, *Operations Management: Decision Making in the Operations Function*, McGraw-Hill, New York, NY, 1993.

M. J. Scott and E. K. Antonsson, "Arrow's Theorem and Engineering Decision Making," *Research in Engineering Design*, *11*, 218–228, 1999.

J. J. Shah, "Experimental Investigation of Progressive Idea Generation Techniques in Engineering Design," *Proceedings of the 1998 ASME Design Theory and Methodology Conference*, American Society of Mechanical Engineers, New York, NY, 1998.

J. E. Shigley, C. R. Mischke, and R. G. Budynas, *Mechanical Engineering Design*, 7th ed., McGraw Hill, New York, 2006.

H. A. Simon, "Style in Design," in C. M. Eastman (Editor), *Spatial Synthesis in Computer-Aided Building Design*, Applied Science Publishers, London, England, 1975.

H. A. Simon, *The Sciences of the Artificial*, 2nd ed., MIT Press, Cambridge, MA, 1981.

A. Slocum, *FUNdaMENTALS of Design*, available gratis at <http://web.mit.edu/2.75/resources/FUNdaMENTALS.html>, last accessed 25 August 2012.

L. Stauffer, *An Empirical Study on the Process of Mechanical Design*, Thesis, Department of Mechanical Engineering, Oregon State University, Corvallis, OR, 1987.

L. Stauffer, D. G. Ullman, and T. G. Dietterich, "Protocol Analysis of Mechanical Engineering Design," *Proceedings of the 1987 International Conference on Engineering Design*, Boston, MA, 1987.

G. Stevens, *The Reasoning Architect: Mathematics and Science in Design*, McGraw-Hill, New York, NY, 1990.

S. Stevenson and S. Whitmore, *Strategies for Engineering Communication*, John Wiley & Sons, Inc., New York, 2002.

G. Stiny and J. Gips, *Algorithmic Aesthetics*, University of California Press, Berkeley, CA, 1978.

N. P. Suh, *Axiomatic Design: Advances and Applications*, Oxford University Press, Oxford, England, 2001.

N. P. Suh, *The Principles of Design*, Oxford University Press, Oxford, England, 1990.

M. C. Thomsett, *The Little Black Book of Project Management*, American Management Association, New York, NY, 1990.

C. Tong and D. Sriram (Editors), *Artificial Intelligence in Engineering Design, Volume I: Design Representation and Models of Routine Design*, Academic Press, Boston, MA, 1992a.

C. Tong and D. Sriram (Editors), *Artificial Intelligence in Engineering Design, Vol. II: Models of Innovative Design, Reasoning about Physical Systems, and Reasoning about Geometry*, Academic Press, Boston, MA, 1992b.

C. Tong and D. Sriram (Editors), *Artificial Intelligence in Engineering Design, Vol. III: Knowledge Acquisition, Commercial Applications and Integrated Environments*, Academic Press, Boston, MA, 1992c.

B. W. Tuckman, "Developmental Sequences in Small Groups," *Psychological Bulletin*, *63*, 384–399, 1965.

E. R. Tufte, *The Visual Display of Quantitative Information*, Graphics Press, Cheshire, CT, 2001.

K. Turabian, *A Manual for Writers of Term Papers, Theses, and Dissertations*, University of Chicago Press, Chicago, 1996.

D. G. Ullman, "A Taxonomy for Mechanical Design," *Research in Engineering Design*, *3*, 1992.

D. G. Ullman, *The Mechanical Design Process*, 2nd ed., McGraw-Hill, New York, NY, 1997.

D. G. Ullman and T. G. Dietterich, "Toward Expert CAD," *Computers in Mechanical Engineering*, *6* (3), 56–70, 1987.

D. G. Ullman, T. G. Dietterich, and L. Stauffer, "A Model of the Mechanical Design Process Based on Empirical Data," *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, *2* (1), 1988.

D. G. Ullman, S. Wood and D. Craig, "The Importance of Drawing in the Mechanical Design Process," *Computers and Graphics*, *14* (2), 1990.

K. T. Ulrich and S. D. Eppinger, *Product Design and Development*, 2nd ed., McGraw-Hill, New York, NY, 2000.

G. N. Vanderplaats, *Numerical Optimization Techniques for Engineering Design*, McGraw-Hill, New York, NY, 1984.

VDI, *VDI-2221: Systematic Approach to the Design of Technical Systems and Products*, Verein Deutscher Ingenieure, VDI-Verlag, Translation of the German Edition 11/1986, 1987.

C. E. Wales, R. A. Stager, and T. R. Long, *Guided Engineering Design: Project Book*, West Publishing Company, St. Paul, MN, 1974.

J. Walton, *Engineering Design: From Art to Practice*, West Publishing, St. Paul, MN, 1991.

T. Wiedmann and J. Minx, "A Definition of 'Carbon Footprint,'" in C. C. Pertsova (Editor), *Ecological Economics Research Trends*, Chapter 1, pp. 1–11, Nova Science Publishers, Hauppauge, NY, 2007.

D. J. Wilde, *Globally Optimal Design*, John Wiley & Sons, Inc., New York, 1978.

B. A. Wilson, *Design Dimensioning and Tolerancing*, Goodheart-Wilcox, Tinley Park, IL, 2005.

L. Winner, "Engineering Ethics and Political Imagination," in P. T. Durbin (Editor), *Philosophy of Technology II: Broad and Narrow Interpretations*, Kluwer Academic Publishers, Dordrecht, Netherlands, 1990.

T. T. Woodson, *Introduction to Engineering Design*, McGraw-Hill, New York, NY, 1966.

R. N. Wright, S. J. Fenves, and J. R. Harris, *Modeling of Standards: Technical Aids for Their Formulation, Expression and Use*, National Bureau of Standards, Washington, D.C., March 1980.

C. Zener, *Engineering Design by Geometric Programming*, Wiley-Interscience, New York, NY, 1971.

C. Zozaya-Gorostiza, C. Hendrickson, and D. R. Rehak, *Knowledge-Based Process Planning for Construction and Manufacturing*, Academic Press, Boston, MA, 1989.

# INDEX