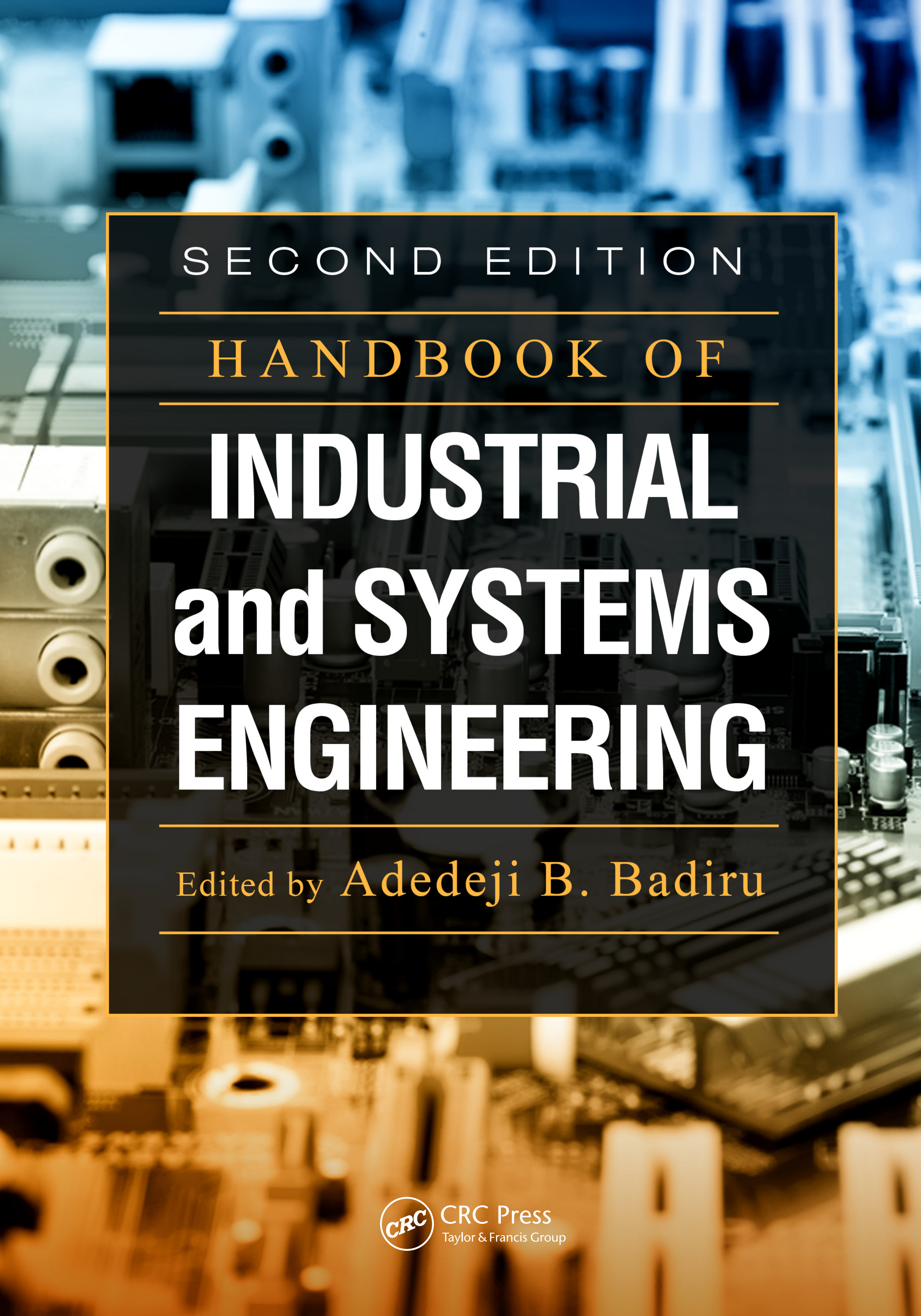
SECOND EDITION
HANDBOOK OF
INDUSTRIAL and SYSTEMS ENGINEERING
Edited by Adedeji B. Badiru
CRC Press
Taylor & Francis Group

SECOND EDITION

HANDBOOK OF

INDUSTRIAL and SYSTEMS ENGINEERING

Industrial Innovation Series

Series Editor

Adedeji B. Badiru

Department of Systems and Engineering Management
Air Force Institute of Technology (AFIT) – Dayton, Ohio

PUBLISHED TITLES

Carbon Footprint Analysis: Concepts, Methods, Implementation, and Case Studies, *Matthew John Franchetti & Defne Apul*

Computational Economic Analysis for Engineering and Industry, *Adedeji B. Badiru & Olufemi A. Omitaomu*

Conveyors: Applications, Selection, and Integration, *Patrick M. McGuire*

Global Engineering: Design, Decision Making, and Communication, *Carlos Acosta, V. Jorge Leon, Charles Conrad, & Cesar O. Malave*

Handbook of Emergency Response: A Human Factors and Systems Engineering Approach, *Adedeji B. Badiru & LeeAnn Racz*

Handbook of Industrial Engineering Equations, Formulas, and Calculations, *Adedeji B. Badiru & Olufemi A. Omitaomu*

Handbook of Industrial and Systems Engineering, Second Edition *Adedeji B. Badiru*

Handbook of Military Industrial Engineering, *Adedeji B.Badiru & Marlin U. Thomas*

Industrial Control Systems: Mathematical and Statistical Models and Techniques, *Adedeji B. Badiru, Oye Ibidapo-Obe, & Babatunde J. Ayeni*

Industrial Project Management: Concepts, Tools, and Techniques, *Adedeji B. Badiru, Abidemi Badiru, & Adetokunboh Badiru*

Inventory Management: Non-Classical Views, *Mohamad Y. Jaber*

Kansei Engineering - 2 volume set

- Innovations of Kansei Engineering, *Mitsuo Nagamachi & Anitawati Mohd Lokman*
- Kansei/Affective Engineering, *Mitsuo Nagamachi*

Knowledge Discovery from Sensor Data, *Auroop R. Ganguly, João Gama, Olufemi A. Omitaomu, Mohamed Medhat Gaber, & Ranga Raju Vatsavai*

Learning Curves: Theory, Models, and Applications, *Mohamad Y. Jaber*

Modern Construction: Lean Project Delivery and Integrated Practices, *Lincoln Harding Forbes & Syed M. Ahmed*

Moving from Project Management to Project Leadership: A Practical Guide to Leading Groups, *R. Camper Bull*

Project Management: Systems, Principles, and Applications, *Adedeji B. Badiru*

Project Management for the Oil and Gas Industry: A World System Approach, *Adedeji B. Badiru & Samuel O. Osisanya*

Quality Management in Construction Projects, *Abdul Razzak Rumane*

Quality Tools for Managing Construction Projects, *Abdul Razzak Rumane*

Social Responsibility: Failure Mode Effects and Analysis, *Holly Alison Duckworth & Rosemond Ann Moore*

Statistical Techniques for Project Control, *Adedeji B. Badiru & Tina Agustiady*

STEP Project Management: Guide for Science, Technology, and Engineering Projects, *Adedeji B. Badiru*

Systems Thinking: Coping with 21st Century Problems, *John Turner Boardman & Brian J. Sauser*

Techonomics: The Theory of Industrial Evolution, *H. Lee Martin*

Triple C Model of Project Management: Communication, Cooperation, Coordination, *Adedeji B. Badiru*

FORTHCOMING TITLES

Communication for Continuous Improvement Projects, *Tina Agustiady & Adedeji B. Badiru*

Cellular Manufacturing: Mitigating Risk and Uncertainty, *John X. Wang*

Essentials of Engineering Leadership and Innovation, *Pamela McCauley-Bush & Lesia L. Crumpton-Young*

Sustainability: Utilizing Lean Six Sigma Techniques, *Tina Agustiady & Adedeji B. Badiru*

Technology Transfer and Commercialization of Environmental Remediation Technology, *Mark N. Goltz*

SECOND EDITION

HANDBOOK OF

INDUSTRIAL and SYSTEMS ENGINEERING

Edited by Adedeji B. Badiru

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Version Date: 20130809

International Standard Book Number-13: 978-1-4665-1505-5 (eBook - PDF)

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Dedication

Dedicated to Blake Badiru, who will carry on the intellectual legacy.

Contents

Foreword

Handbook of Industrial and Systems Engineering

While any handbook on a major topic such as industrial engineering and systems will have inclusions and omissions, this handbook contains coverage on virtually all areas of industrial engineering viewed from a systems perspective. More than 40 authors have contributed state-of-the-art coverage of the most important modern industrial and systems engineering topics. Most of the authors are the "rising stars" of our field that are at the "cutting edge" of theory and practice. This handbook will be invaluable as a reference guide, classroom text, or support for the professional.

The premise of this new handbook is to incorporate more of the systems engineering aspects of industrial engineering than the two existing handbooks in our field do. The material is presented by the newer authors of our field who have fresh (and possibly revolutionary) ideas. The objective of this handbook is to provide students, researchers, and practitioners with a comprehensive yet concise, easy-to-use guide to a wide range of industrial and systems tools and techniques. The editor's attributes for this handbook, which have been successfully met, in my opinion, include the following:

1. To provide a one-stop reference for industrial and systems engineering
2. To use a comprehensive yet concise format
3. To have an up-to-date treatment of topics
4. To introduce new technology for industrial and systems engineering
5. To use the systems integration approach
6. To provide coverage of information engineering
7. To have diversity of contributions from both industry and academia
8. To provide up-to-date material for teaching, research, and practice

The editor, Dr. Adedeji B. Badiru, is a well-recognized and respected authority and leader in the fields of industrial and systems engineering with numerous academic and professional publications. I know and respect Dr. Badiru and I am proud to have been his dissertation advisor at the University of Central Florida.

Gary E. Whitehouse
Provost Emeritus
Distinguished University Professor
University of Central Florida
Orlando, Florida

Preface

The second edition of the *Handbook of Industrial and Systems Engineering* is an updated collation of the body of knowledge of industrial and systems engineering. The handbook has been substantively expanded from the 36 seminal chapters in the first edition to 56 landmark chapters in the second edition. In addition to the 20 new chapters, 11 of the chapters in the first edition have been updated with new materials.

As with the first edition, the objective of the handbook is to provide students, researchers, and practitioners with a comprehensive and easy access to a wide range of industrial and systems engineering tools and techniques in a concise format. There is a growing need for a handbook on the diverse and versatile field of industrial and systems engineering. The handbook has the following attributes:

1. One-stop reference for industrial and systems engineering
2. Comprehensive and yet concise
3. Up-to-date treatment of topics
4. Introduction of new technology for industrial and systems engineering
5. Systems integration approach
6. Coverage of information engineering
7. Diversification of contributions
8. Up-to-date information for teaching, research, and practice

The handbook fills the gap that exists between the traditional and modern practice of industrial and systems engineering. The overall organization of the book is integrative with respect to quantitative models, qualitative principles, and computer techniques. Where applicable, the handbook encourages a project model for end-of-chapter exercises rather than typical textbook exercises. This is to provide open-ended problem exercises for readers. Most systems issues are open-ended challenges that are best handled from an integrated project perspective.

Part I of the book covers general introduction with specific reference to the origin of industrial engineering and the ties to the Industrial Revolution. Part II covers the fundamentals of industrial engineering. Part III covers the fundamentals of systems engineering. Part IV contains chapters on manufacturing, production systems, and ergonomics. Part V presents chapters on economic and financial analysis. Part VI covers management, information engineering, and decision making. A new Part VII has been added to this second edition to cover safety, reliability, and quality. Also, a new and distinct Part VIII is included in this second edition to cover operations research, queuing, logistics, and scheduling. The appendix has been expanded in this second edition to include two parts. Appendix A contains conversion factors, whereas Appendix B contains engineering, systems, and statistical formulae.

The premise of the handbook remains to expand the breadth and depth of coverage beyond the traditional handbooks on industrial engineering. I strongly believe this pursuit has been fulfilled.

Adedeji B. Badiru
Beavercreek, Ohio

Acknowledgments

Compiling a second edition of an edited handbook requires far more cooperation and coordination than an initial edition. Pulling off the endeavor successfully is an indication of the collaborative spirit of the chapter contributors. For this, I thank all those who subjected themselves to the frequent harassment of a pesky editor. The eventual quality of this handbook could not have been possible without your intellectual contributions.

I thank my proverbial second half, Iswat, as usual, for sustaining the home front while I, once again, took the path of a lengthy writing engagement, despite the several broken promises of "no more" book writing projects. I thank my colleagues, who tolerated and endured my frequent intellectual sequestration while I focused on the physical and mental demands of compiling this expanded handbook. I thank my former professors and unflinching mentors, Dr. Sid Gilbreath and Dr. Gary Whitehouse, for continuing to be sources of inspiration for my intellectual pursuits.

Several individuals participated directly in putting together this edited work. I could not have done it without the administrative, editorial, typing, organizing, and/or copyediting support of my Air Force Institute of Technology graduate students, Kalyn Tung, Robert Poisson, Jacob Petter, and Kelsey Smith. Kalyn's contribution is particularly noteworthy for providing editorial finesse for the initial processing and review of the raw chapters.

I thank all the editorial and production staff members of CRC Press/Taylor & Francis for their expertly expedient manner of handling the publication of this book. I particularly thank Cindy Renee Carelli, senior acquisitions editor, for gently, but firmly, broaching the idea of doing a second edition of this handbook. Although she called it an "invitation," I took it as an order. Thank goodness, it worked out.

Ciao to everyone, until the next edition.

Editor

Adedeji Badiru is professor and head of systems and engineering management at the Air Force Institute of Technology. He was previously professor and department head of industrial and information engineering at the University of Tennessee in Knoxville. Before that, he was professor of industrial engineering and dean of University College at the University of Oklahoma. He is a registered professional engineer (PE), a certified Project Management Professional (PMP), a fellow of the Institute of Industrial Engineers, and a fellow of the Nigerian Academy of Engineering. He holds BS in industrial engineering, MS in mathematics, and MS in industrial engineering from Tennessee Technological University, and PhD in industrial engineering from the University of Central Florida. His areas of interest include mathematical modeling, project modeling and analysis, economic analysis, systems engineering, and efficiency/productivity analysis and improvement. He is the author of over two dozen books and scores of book chapters, refereed journal articles, conference proceedings, and presentations. He also has over two dozen published magazine articles, editorials, and periodicals. He is a member of several professional associations and several scholastic honor societies. Prof. Badiru has won several awards for his teaching, research, and professional accomplishments. He is the recipient of the 2009 Dayton Affiliate Society Council Award for Outstanding Scientists and Engineers in the Education category with a commendation from the 128th Senate of Ohio. He also won the 2010 IIE/Joint Publishers Book-of-the-Year Award for co-editing *The Handbook of Military Industrial Engineering*. He also won the 2010 American Society for Engineering Education John Imhoff Award for his global contributions to industrial engineering education, the 2011 Federal Employee of the Year Award in the Managerial Category from the International Public Management Association, Wright Patterson Air Force Base, the 2012 Distinguished Engineering Alum Award from the University of Central Florida, and the 2012 Medallion Award from the Institute of Industrial Engineers for his global contributions in the advancement of the profession. He has served as a consultant to several organizations worldwide. He holds a leadership certificate from the University Tennessee Leadership Institute. He is also a program evaluator for ABET. Prof. Badiru's professional accomplishments are coupled with his passion for writing about everyday events, interpersonal issues, and socially responsible service to the community. Outside of the academic realm, he writes self-help books, motivational poems, editorials, and newspaper commentaries, as well as engaging in paintings and crafts. He also manages an STEM-and-sports education web site. He is also the founder of the Association of Military Industrial Engineers (AMIE).

Contributors

Navin Acharya
Pennsylvania State University
University Park, Pennsylvania

H. Aiyappan
The University of Alabama at Tuscaloosa
Tuscaloosa, Alabama

Mehmet Aktan
Atatürk University
Erzurum, Turkey

Michihiro Amagasa
Hokkai Gakuen University
Hokkaido, Japan

Sunday Asaolu
University of Lagos
Lagos, Nigeria

M. Affan Badar
Indiana State University
Terra Haute, Indiana

Adedeji B. Badiru
Air Force Institute of Technology
Wright-Patterson Air Force Base, Ohio

Amarnath Banerjee
Texas A&M University
College Station, Texas

Jonathan F. Bard
Graduate Program in Operation,
Research and Industrial Engineering
University of Texas
Austin, Texas

David Ben-Arieh
Kansas State University
Manhattan, Kansas

Richard E. Billo
University of Texas at Arlington
Arlington, Texas

Jennifer Blackhurst
Iowa State University
Ames, Iowa

Yi-Yun Chang
National Taiwan University of Science and Technology
Taipei, Taiwan

Zhifeng Chen
Discover Financial Services
Riverwoods, Illinois

Altyn Clark
Transformation Systems, Inc.
Arlington, Virginia

Garry D. Coleman
Transformation Systems, Inc.
Arlington, Virginia

John M. Colombi
Air Force Institute of Technology
Wright-Patterson Air Force Base, Ohio

Lesia L. Crumpton-Young
University of Central Florida
Orlando, Florida

Taren Daigle
Louisiana State University
Baton Rouge, Louisiana

Ashok Darisipudi
Louisiana State University
Baton Rouge, Louisiana

Gail W. DePuy
University of Louisville
Louisville, Kentucky

Ike C. Ehie
Kansas State University
Manhattan, Kansas

Matthew Elam
Private Consultant
Texas

John V. Farr
Stevens Institute of Technology
Hoboken, New Jersey

Bobbie Leon Foote
U.S. Military Academy
West Point, New York

Lincoln H. Forbes
Department of Facilities Design and Standards
Miami-Dade County Public Schools
Miami, Florida

Ronald Giachetti
Department of Systems Engineering
Naval Postgraduate School
Monterey, California

Craig M. Harvey
Louisiana State University
Baton Rouge, Louisiana

Alan R. Heminger
Air Force Institute of Technology
Wright-Patterson Air Force Base, Ohio

O. Ibidapo-Obe
Federal University
Ebonyi State, Nigeria

Mohamad Y. Jaber
Ryerson University
Toronto, Canada

Madhu Jain
Banasthali University
Banasthali, India

Eugene Rex Jalao
Arizona State University
Phoenix, Arizona

Steve Jarvis
Jarvis Bagshaw Ltd.
Milton Keynes, United Kingdom
and
Cranfield University
Cranfield, United Kingdom

Wei Jiang
Stevens Institute of Technology
Hoboken, New Jersey

Andrew L. Johnson
Texas A&M University
College Station, Texas

Victoria Jordan
Quality Engineering and Clinical Operations Informatics
University of Texas
MD Anderson Cancer Center
Houston, Texas

Manjunath Kamath
School of Industrial Engineering and Management
Oklahoma State University
Stillwater, Oklahoma

Kyoung-Yun Kim
Wayne State University
Detroit, Michigan

Seong Kim
Pennsylvania State University
University Park, Pennsylvania

Dongjoon Kong
Boeing Company
Seattle, Washington

Chia-Yen Lee
National Cheng Kung University
Tainan City, Taiwan

Tzong-Ru (Jiun-Shen) Lee
National Chung-Hsing University
Taichung City, Taiwan

Gino Lim
Department of Industrial Engineering
University of Houston
Houston, Texas

Pamela R. McCauley-Bell
University of Central Florida
Orlando, Florida

Maethee Mekaroonreung
Texas A&M University
College Station, Texas

Michael E. Miller
Air Force Institute of Technology
Wright-Patterson Air Force Base, Ohio

Arezou Mobasher
Department of Industrial Engineering
University of Houston
Houston, Texas

Reinaldo J. Moraga
Northern Illinois University De Kalb
DeKalb, Illinois

Johnathan Mun
Naval Postgraduate School
Monterey, California

Larry Nabatilan
Louisiana State University
Baton Rouge, Louisiana

David A. Nembhard
Pennsylvania State University
University Park, Pennsylvania

Harriet Black Nembhard
Pennsylvania State University
University Park, Pennsylvania

Bartholomew O. Nnaji
Geometric Power
Abuja, Nigeria

Bryan A. Norman
University of Pittsburgh
Pittsburgh, Pennsylvania

S.A. Oke
University of Lagos
Lagos, Nigeria

Albert Olagbemiro
US Air Force Reserve

Olufemi A. Omitaomu
Oak Ridge National Laboratory
Oak Ridge, Tennessee

Brian Peacock
National University of Singapore
Singapore
and
SIM University Singapore
Singapore

Brandon Pope
Purdue University
Lafayette, Indiana

J. David Porter
Oregon State University
Corvallis, Oregon

G.N. Purohit
Banasthali University
Banasthali, India

Shivakumar Raman
University of Oklahoma
Norman, Oklahoma

Shinu Rani
Banasthali University
Banasthali, India

Paul S. Ray
The University of Alabama at Tuscaloosa
Tuscaloosa, Alabama

Marc Resnick
Bentley University
Waltham, Massachusetts

Ling Rothrock
Pennsylvania State University
University Park, Pennsylvania

M.D. Sarder
University of Southern Mississippi
Hattiesburg, Mississippi

Rupy Sawhney
University of Tennessee
Knoxville, Tennessee

Dujuan B. Sevillian
Cranfield University
Cranfield, United Kingdom
and
The Boeing Company
Bedforshire, United Kingdom

Man-Tak Shing
Naval Postgraduate School
Monterey, California

Arif Sirinterlikci
Robert Morris University
Coraopolis, Pennsylvania

Yi-shiang Shiou
National Chung-Hsing University
Taichung City, Taiwan

Som Soni
AdTech Systems, Inc. Dayton, Ohio
Oklahoma State University
Stillwater, Oklahoma

Sandeep Srivathsan
Nanyang Business School
Nanyang Technological University
Singapore

Evangelos Triantaphyllou
Louisiana State University
Baton Rouge, Louisiana

Anthony P. Tvaryanas
711th Human Performance Wing,
US Air Force

Ricardo Valerdi
University of Arizona
Tucson, Arizona

Aashish Wadke
The University of Oklahoma
Norman, Oklahoma

Xiaoting Wang
Louisiana State University
Baton Rouge, Louisiana

Yan Wang
University of Pittsburgh
Pittsburgh, Pennsylvania

Gary E. Whitehouse
University of Central Florida
Orlando, Florida

Teresa Wu
Arizona State University
Phoenix, Arizona

Nong Ye
Arizona State University
Phoenix, Arizona

Deborah Young-Corbett
Virginia Tech
Blacksburg, Virginia

part one

General introduction

chapter one

General introduction

Adedeji B. Badiru

Contents

> "Think like an IE, act like an IE."
> Adedeji Badiru's motto for the practice of industrial engineering

Have you ever wondered …

- How a product can be designed to fit people, rather than forcing people to accommodate the product?
- How merchandise layouts can be designed to maximize the profit of a retail store?
- How hospitals can improve patient care while lowering cost?
- How paper companies manage their forests (paper-making raw material) to both increase profits and still ensure long-term availability of trees?
- How the work environment can be designed to enhance comfort and safety while increasing productivity?
- How a fast-food restaurant knows how many and which kinds of burgers to have ready for the lunch-break rush?
- How new car designs can be tested before a prototype is ever built?
- How space exploration can be coordinated to link both management and technical requirements?
- How a military multi-pronged attack can be organized to sustain the supply lines?

Industrial engineers, with a systems thinking approach, help answer and solve all these questions. Industrial engineering thrives on systems perspectives just as systems thrive on industrial engineering approaches. One cannot treat topics of industrial engineering effectively without recognizing systems perspectives and vice versa. Thus, it makes sense to have a handbook that integrates industrial and systems engineering (ISE) principles. A generic definition of an industrial engineering, adopted by the Institute of Industrial Engineers states

> Industrial Engineer—one who is concerned with the design, installation, and improvement of integrated systems of people, materials,

> information, equipment, and energy by drawing upon specialized knowledge and skills in the mathematical, physical, and social sciences, together with the principles and methods of engineering analysis and design to specify, predict, and evaluate the results to be obtained from such systems.

The above definition embodies the various aspects of what an industrial engineer does. Although some practitioners find the definition to be too convoluted, it, nonetheless, describes an industrial engineer. As can be seen, the profession is very versatile, flexible, and diverse. It can also be seen from the definition that a systems orientation permeates the work of industrial engineers. Some of the major functions of industrial engineers involve the following:

- Design integrated systems of people, technology, processes, and methods
- Develop performance modeling, measurement, and evaluation for systems
- Develop and maintain quality standards for industry and business
- Apply production principles to pursue improvements in service organizations
- Incorporate technology effectively into work processes
- Develop cost mitigation, avoidance, or containment strategies
- Improve overall productivity of integrated systems of people, materials, and processes
- Recognize and incorporate factors affecting performance of a composite system
- Plan, organize, schedule, and control production and service projects
- Organize teams to improve efficiency and effectiveness of an organization
- Install technology to facilitate work flow
- Enhance information flow to facilitate smooth operations of systems
- Coordinate materials and equipment for effective systems performance

1.1 *What is industrial engineering?*

> Industrial engineering makes systems function better together with less waste, better quality, and fewer resources.
>
> **Susan Blake**
> *Industrial Engineer*
> *Tinker Air Force Base, 2011*

The goal of every organization is to eliminate waste. Thus, the above definition is aptly relevant for everyone. Industrial engineering can be described as the practical application of the combination of engineering fields together with the principles of scientific management. It is the engineering of work processes and the application of engineering methods, practices, and knowledge to production and service enterprises. Industrial engineering places a strong emphasis on an understanding of workers and their needs in order to increase and improve production and service activities. Industrial engineering activities and techniques include the following:

1. Designing jobs (determining the most economic way to perform work)
2. Setting performance standards and benchmarks for quality, quantity, and cost
3. Designing and installing facilities

1.2 What is systems engineering?

Systems engineering involves a recognition, appreciation, and integration of all aspects of an organization or a facility. A system is defined as a collection of interrelated elements working together in synergy to produce a composite output that is greater than the sum of the individual outputs of the components. A systems view of a process facilitates a comprehensive inclusion of all the factors involved in the process.

1.3 Ties to the Industrial Revolution

Industrial engineering has a proud heritage with a link that can be traced back to the Industrial Revolution. Although the practice of industrial engineering has been in existence for centuries, the work of Frederick Taylor in the early 20th century was the first formal emergence of the profession. It has been referred to with different names and connotations. Scientific management was one of the original names used to describe what industrial engineers do.

Industry, the root of the profession's name, clearly explains what the profession is about. The dictionary defines industry generally as the ability to produce and deliver goods and services. The "industry" in industrial engineering can be viewed as the application of skills and cleverness to achieve work objectives. This relates to how human effort is harnessed innovatively to carry out work. Thus, any activity can be defined as "industry" because it generates a product—be it a service or a physical product. A systems view of industrial engineering encompasses all the details and aspects necessary for applying skills and cleverness to produce work efficiently. However, the academic curriculum of industrial engineering must change, evolve, and adapt to the changing systems environment of the profession.

It is widely recognized that the occupational discipline that has contributed the most to the development of modern society is *engineering*, through its various segments of focus. Engineers design and build infrastructures that sustain the society. These include roads, residential and commercial buildings, bridges, canals, tunnels, communication systems, healthcare facilities, schools, habitats, transportation systems, and factories. Across all of these, the industrial engineering process of systems integration facilitates the success of the infrastructures. In this sense, the scope of ISE steps through the levels of activity, task, job, project, program, process, system, enterprise, and society. This handbook of ISE presents essential tools for the levels embodied by this hierarchy of functions. From the age of horse-drawn carriages and steam engines to the present age of intelligent automobiles and aircraft, the impacts of ISE cannot be mistaken, even though the contributions may not be recognized in the context of the ISE disciplinary identification.

It is essential to recognize the alliance between "industry" and industrial engineering as the core basis for the profession. The profession has gone off on too many different tangents over the years. Hence, it has witnessed the emergence of industrial engineering professionals who claim sole allegiance to some narrow line of practice, focus, or specialization rather than the core profession itself. Industry is the original basis of industrial engineering and it should be preserved as the core focus, which should be supported by the different areas of specialization. While it is essential that we extend the tentacles of industrial engineering to other domains, it should be realized that overdivergence of practice will not sustain the profession. The continuing fragmentation of industrial engineering is a major reason to compile a handbook such as this. A fragmented profession cannot survive for long. The incorporation of systems can help bind everything together.

Notable industrial developments that fall under the purview of the practice of industrial engineering range from the invention of the typewriter to the invention of the automobile. Some examples are presented below.

1.4 Typewriter history

Writing is a basic means of communicating and preserving records. It is one of the most basic accomplishments of society. The course of history might have taken a different path if early writing instruments had not been invented at the time that they were. Below is the chronological history of the typewriter:

1714: Henry Mill obtained British patent for a writing machine.
1833: Xavier Progin created a machine that uses separate levers for each letter.
1843: American inventor, Charles Grover Thurber, developed a machine that moves paper horizontally to produce spacing between lines.
1873: E. Remington & Sons of Ilion, New York, manufacturers of rifles and sewing machines, developed a typewriter patented by Carlos Glidden, Samuel W. Soule, and Christopher Latham Sholes, who designed the modern keyboard. This class of typewriters wrote in only uppercase letters but contained most of the characters on the modern machines.
1912: Portable typewriters were first introduced.
1925: Electric typewriters became popular. This made typeface to be more uniform. International Business Machines Corporation (IBM) was a major distributor for this product.

In each case of product development, engineers demonstrate the ability to design, develop, manufacture, implement, and improve integrated systems that include people, materials, information, equipment, energy, and other resources. Thus, product development must include an in-depth understanding of appropriate analytical, computational, experimental, implementation, and management processes.

1.5 Heritage of industrial and systems engineering

Going further back in history, several developments helped form the foundation for what later became known as industrial engineering. In America, George Washington was said to have been fascinated by the design of farm implements on his farm in Mt. Vernon. He has an English manufacturer send him a plow built to his specifications that included a mold on which to form new irons when old ones were worn out, or would need repairs. This can be described as one of the early attempts to create a process of achieving a system of interchangeable parts. Thomas Jefferson invented a wooden mold board, which, when fastened to a plow, minimized the force required to pull the plow at various working depths. This is an example of early agricultural industry innovation. Jefferson also invented a device that allowed a farmer to seed four rows at a time. In pursuit of higher productivity, he invented a horse-drawn threshing machine that did the work of 10 men.

Meanwhile in Europe, the Industrial Revolution was occurring at a rapid pace. Productivity growth, through reductions in manpower, marked the technological innovations of 1769–1800 Europe. Sir Richard Arkwright developed a practical code of factory discipline. In their foundry, Matthew Boulton and James Watt developed a complete and integrated engineering plant to manufacture steam engines. They developed extensive methods of market research, forecasting, plant location planning, machine layout, work flow, machine operating standards, standardization of product components, worker training, division of labor, work study, and other creative approaches to increasing productivity. Charles Babbage, who is credited with the first idea of a computer, documented ideas on scientific methods of managing industry in his book entitled *On the Economy of Machinery and Manufacturers*, which was first published in 1832. The book contained ideas on division of labor, paying less for less important tasks, organization charts, and labor relations. These were all forerunners of modern industrial engineering.

Back in America, several efforts emerged to form the future of the industrial engineering profession. Eli Whitney used mass production techniques to produce muskets for the US Army. In 1798, Whitney developed the idea of having machines make each musket part so that it could be interchangeable with other similar parts. By 1850, the principle of interchangeable parts was widely adopted. It eventually became the basis for modern mass production for assembly lines. It is believed that Eli Whitney's principle of interchangeable parts contributed significantly to the Union victory during the US Civil War.

The management attempts to improve productivity before 1880 did not consider the human element as an intrinsic factor. However, from 1880 through the first quarter of the 20th century, the works of Frederick W. Taylor, Frank and Lillian Gilbreth, and Henry L. Gantt created a long-lasting impact on productivity growth through consideration of the worker and his or her environment.

Frederick Winslow Taylor (1856–1915) was born in the Germantown section of Philadelphia to a well-to-do family. At the age of 18, he entered the labor force, having abandoned his admission to Harvard University because of an impaired vision. He became an apprentice machinist and pattern maker in a local machine shop. In 1878, when he was 22, he went to work at the Midvale Steel Works. The economy was in a depressed state at the time. Frederick was employed as a laborer. His superior intellect was very quickly recognized. He was soon advanced to the positions of time clerk, journeyman, lathe operator, gang boss, and foreman of the machine shop. By the age of 31, he was made chief engineer of the company. He attended night school and earned a degree in mechanical engineering in 1883 from Stevens Institute. As a work leader, Taylor faced the following common questions:

"Which is the best way to do this job?"
"What should constitute a day's work?"

These are still questions faced by industrial and systems engineers of today. Taylor set about the task of finding the proper method for doing a given piece of work, instructing the worker in following the method, maintaining standard conditions surrounding the work so that the task could be properly accomplished, and setting a definite time standard and payment of extra wages for doing the task as specified. Taylor later documented his industry management techniques in his book entitled *The Principles of Scientific Management.*

The work of Frank and Lillian Gilbreth coincided with the work of Frederick Taylor. In 1895, on his first day on the job as a bricklayer, Frank Gilbreth noticed that the worker assigned to teach him how to lay brick did his work three different ways. The bricklayer was insulted when Frank tried to tell him of his work inconsistencies—when training someone on the job, when performing the job himself, and when speeding up. Frank thought it was essential to find one best way to do the work. Many of Frank Gilbreth's ideas were similar to Taylor's ideas. However, Gilbreth outlined procedures for analyzing each step of the work flow. Gilbreth made it possible to apply science more precisely in the analysis and design of the workplace. Developing *therbligs*, which is Gilbreth spelled backward, as elemental predetermined time units, Frank and Lillian Gilbreth were able to analyze the motions of a worker in performing most factory operations in a maximum of 18 steps. Working as a team, they developed techniques that later became known as work design, methods improvement, work simplification, value engineering, and optimization. Lillian (1878–1972) brought to the engineering profession the concern for human relations. The foundation for establishing the profession of industrial engineering was originated by Frederick Taylor and Frank and Lillian Gilbreth.

Henry Gantt's work advanced the management movement from an industrial management perspective. He expanded the scope of managing industrial operations. His concepts emphasized the unique needs of the worker by recommending the following considerations for managing work:

1. Define his task, after a careful study.
2. Teach him how to do it.
3. Provide an incentive in terms of adequate pay or reduced hours.
4. Provide an incentive to surpass it.

Henry Gantt's major contribution is the Gantt chart, which went beyond the works of Frederick Taylor or the Gilbreths. The Gantt chart related every activity in the plant to the factor of time. This was a revolutionary concept at the time. It led to better production planning control and better production control. This involved visualizing the plant as a whole, like one big system made up of interrelated subsystems. Table 1.1 summarizes the major chronological events marking the origin of ISE. As can be seen from the table, industry has undergone a hierarchical transformation over the past several decades. Figure 1.1 shows how industry has been transformed from one focus level to the next ranging from efficiency of the 1960s to the present-day nanoscience trend. It shows the progression from the classical efficiency focus to the present and future platforms of cyber operations.

In pursuing the applications of ISE, it is essential to make a distinction between the tools, techniques, models, and skills of the profession. *Tools* are the instruments, apparatus, and devices (usually visual or tangible) that are used for accomplishing an objective. *Techniques* are the means, guides, and processes for utilizing tools for accomplishing the objective. A simple and common example is the technique of using a hammer (a tool) to strike a nail to drive the nail into a wooden work piece (objective). A *model* is a bounded series of steps, principles, or procedures for accomplishing a goal. A model applied to one problem can be replicated and reapplied to other similar problems, provided the boundaries of the model fit the scope of the problem at hand. *Skills* are the human-based processes of using tools, techniques, and models to solve a variety of problems. Very important within the skills set of an industrial engineer are interpersonal skills or soft skills. This human-centric attribute of industrial engineering is what sets it apart from other engineering fields. Table 1.2 summarizes examples of tools, techniques, and skills of ISE.

Table 1.1 Major Chronological Events Marking the Origin of Industrial and Systems Engineering

Year	Major publications and events
1440	Venetian ships were reconditioned and refitted on an assembly line.
1474	Venetian senate passed the first patent law and other industrial laws.
1568	Jacques Besson published an illustrated book on iron machinery as replacement for wooden machines.
1622	William Oughtred invented the slide rule.
1722	Rene de Reaunur published the first handbook on iron technology.
1733	John Kay patented the flying shuttle for textile manufacture—a landmark in textile mass production.
1747	Jean Rodolphe Perronet established the first engineering school.
1765	Watt invented the separate condenser, which made the steam engine the power source.
1770	James Hargreaves patented his "Spinning Jenny." Jesse Ramsden devised a practical screw-cutting lathe.
1774	John Wilkinson built the first horizontal boring machine.
1775	Richard Arkwright patented a mechanized mill in which raw cotton is worked into thread.
1776	James Watt built the first successful steam engine, which became a practical power source.
1776	Adam Smith discussed the division of labor in *The Wealth of Nations.*
1785	Edmund Cartwright patented a power loom.
1793	Eli Whitney invented the "cotton gin" to separate cotton from its seeds.
1797	Robert Owen used modern labor and personnel management techniques in a spinning plant in the New Lanark Mills in Manchester, England.
1798	Eli Whitney designed muskets with interchangeable parts.
1801	Joseph Marie Jacquard designed automatic control for pattern-weaving looms using punched cards.
1802	The "Health and Morals Apprentices Act" in Britain aimed at improving standards for young factory workers.
	Marc Isambard Brunel, Samuel Benton, and Henry Maudsey designed an integrated series of 43 machines to mass produce pulley blocks for ships.
1818	The Institution of Civil Engineers was founded in Britain.
1824	The repeal of the Combination Act in Britain legalized trade unions.
1829	Mathematician Charles Babbage designed "analytical engine," a forerunner of the modern digital computer.
1831	Charles Babbage published *On the Economy of Machines and Manufacturers.*
1832	The Sadler Report exposed the exploitation of workers and the brutality practiced within factories.
1833	The Factory Law was enacted in the United Kingdom. The Factory Act regulated British children's working hours.
	A general Trades Union was formed in New York.
1835	Andrew Ure published *Philosophy of Manufacturers*.
	Samuel Morse invented the telegraph.
1845	Friederich Engels published *Condition of the Working Classes in England.*
1847	The Factory Act in Britain reduced the working hours of women and children to 10 hours per day.
	George Stephenson founded the Institution of Mechanical Engineers.

(*continued*)

Table 1.1 (Continued) Major Chronological Events Marking the Origin of Industrial and Systems Engineering

Year	Major publications and events
1856	Henry Bessemer revolutionized the steel industry through a novel design for a converter.
1869	A transcontinental railroad was completed in the United States.
1871	British Trade Unions were legalized by the Act of Parliament.
1876	Alexander Graham Bell invented a usable telephone.
1877	Thomas Edison invented the phonograph.
1878	Frederick W. Taylor joined Midvale Steel Company.
1880	The American Society of Mechanical Engineers (ASME) was organized.
1881	Frederick Taylor began time study experiments.
1885	Frank B. Gilbreth began motion study research.
1886	Henry R. Towne presented the paper, *The Engineer as Economist*.
	The American Federation of Labor (AFL) was organized.
	Vilfredo Pareto published *Course in Political Economy*.
	Charles M. Hall and Paul L. Herault independently invented an inexpensive method of making aluminum.
1888	Nikola Tesla invented the alternating current induction motor, which enabled electricity to take over from steam as the main provider of power for industrial machines.
	Dr. Herman Hollerith invented the electric tabulator machine, the first successful data processing machine.
1890	The Sherman Anti-Trust Act was enacted in the United States.
1892	Gilbreth completed a motion study of bricklaying.
1893	Taylor began work as a consulting engineer.
1895	Taylor presented the paper entitled *A Piece-Rate System* to ASME.
1898	Taylor began time study at Bethlehem Steel.
	Taylor and Maunsel White developed a process for heat-treating high-speed tool steels.
1899	Carl G. Barth invented a slide rule for calculating metal cutting speed as part of the Taylor system of management.
1901	American National Standards were established.
	Yawata Steel began operation in Japan.
1903	Taylor presented the paper entitled *Shop Management* to ASME.
	H.L. Gantt developed the "Gantt Chart."
	Hugo Diemers wrote *Factory Organization and Administration*.
	Ford Motor Company was established.
1904	Harrington Emerson implemented Santa Fe Railroad improvement.
	Thorstein B. Veblen: *The Theory of Business Enterprise.*
1906	Taylor established the metal-cutting theory for machine tools.
	Vilfredo Pareto: *Manual of Political Economy*.
1907	Gilbreth used time study for construction.
1908	Model T Ford was built.
	Pennsylvania State College introduced the first university course in industrial engineering.
1911	Taylor published *The Principles of Scientific Management*.
	Gilbreth published *Motion Study*.
	The Factory Laws were enacted in Japan.

(*continued*)

Table 1.1 (Continued) Major Chronological Events Marking the Origin of Industrial and Systems Engineering

Year	Major publications and events
1912	Harrington Emerson published *The Twelve Principles of Efficiency.*
	Frank and Lillian Gilbreth presented the concept of *therbligs*.
	Yokokawa translated into Japanese Taylor's *Shop Management* and *The Principles of Scientific Management.*
1913	Henry Ford established a plant at Highland Park, Michigan, which utilized the principles of uniformity and interchangeability of parts, and of the moving assembly line by means of conveyor belt.
	Hugo Munstenberg published *Psychology of Industrial Efficiency.*
1914	World War I.
	Clarence B. Thompson edited *Scientific Management*, a collection of articles on Taylor's system of management.
1915	Taylor's system was used at Niigata Engineering's Kamata plant in Japan.
	Robert Hoxie published *Scientific Management and Labour*.
1916	Lillian Gilbreth published *The Psychology of Management.*
	The Taylor Society was established in the United States.
1917	The Gilbreths published *Applied Motion Study.*
	The Society of Industrial Engineers was formed in the United States.
1918	Mary P. Follet published *The New State: Group Organization, the Solution of Popular Government.*
1919	Henry L. Gantt published *Organization for Work.*
1920	Merrick Hathaway presented the paper *Time Study as a Basis for Rate Setting.*
	General Electric established divisional organization.
	Karel Capek: *Rossum's Universal Robots.* This play coined the word "robot."
1921	The Gilbreths introduced process analysis symbols to ASME.
1922	Toyoda Sakiichi's automatic loom was developed.
	Henry Ford published *My Life and Work.*
1924	The Gilbreths announced the results of their micromotion study using *therbligs*.
	Elton Mayo conducted illumination experiments at Western Electric.
1926	Henry Ford published *Today and Tomorrow*.
1927	Elton Mayo and others began a relay-assembly test room study at the Hawthorne plant.
1929	Great Depression.
	The International Scientific Management Conference was held in France.
1930	Hathaway: *Machining and Standard Times.*
	Allan H. Mogensen discussed 11 principles for work simplification in *Work Simplification.*
	Henry Ford published *Moving Forward*.
1931	Dr. Walter Shewhart published *Economic Control of the Quality of Manufactured Product.*
1932	Aldous Huxley published *Brave New World*, the satire that prophesied a horrifying future ruled by industry.
1934	General Electric performed micromotion studies.
1936	The word "automation" was first used by D.S. Harder of General Motors. It was used to signify the use of transfer machines, which carry parts automatically from one machine to the next, thereby linking the tools into an integrated production line.

(*continued*)

Table 1.1 (Continued) Major Chronological Events Marking the Origin of Industrial and Systems Engineering

Year	Major publications and events
1936	Charlie Chaplin produced *Modern Times*, a film showing an assembly line worker driven insane by routine and unrelenting pressure of his job.
1937	Ralph M. Barnes published *Motion and Time Study*.
1941	R.L. Morrow: *Ratio Delay Study*, an article in the *Mechanical Engineering* journal.
	Fritz J. Roethlisberger: *Management and Morale*.
1943	The ASME work standardization committee published a glossary of industrial engineering terms.
1945	Marvin E. Mundel devised "memo-motion" study, a form of work measurement using time-lapse photography.
	Joseph H. Quick devised the work factors (WF) method.
1945	Shigeo Shingo presented the concept of production as a network of processes and operations, and identified lot delays as source of delay between processes, at a technical meeting of the Japan Management Association.
1946	The first all-electronic digital computer ENIAC (Electronic Numerical Integrator and Computer) was built at Pennsylvania University.
	The first fully automatic system of assembly was applied at the Ford Motor Plant.
1947	American mathematician, Norbert Wiener: *Cybernetics*.
1948	H.B. Maynard and others introduced the methods time measurement (MTM) method.
	Larry T. Miles developed value analysis (VA) at General Electric.
	Shigeo Shingo announced process-based machine layout.
	The American Institute of Industrial Engineers was formed.
1950	Marvin E. Mundel: *Motion and Time Study, Improving Productivity.*
1951	Inductive statistical quality control was introduced to Japan from the United States.
1952	A role and sampling study of industrial engineering was conducted at ASME.
1953	B.F. Skinner: *Science of Human Behaviour*.
1956	A new definition of industrial engineering was presented at the American Institute of Industrial Engineering convention.
1957	Chris Argyris: *Personality and Organization.*
	Herbert A. Simon: *Organizations.*
	R.L. Morrow: *Motion and Time Study.*
1957	Shigeo Shingo introduced scientific thinking mechanism (STM) for improvements.
	The Treaty of Rome established the European Economic Community.
1960	Douglas M. McGregor: *The Human Side of Enterprise.*
1961	Rensis Lickert: *New Patterns of Management*.
1961	Shigeo Shingo devised ZQC (source inspection and poka-yoke systems).
1961	Texas Instruments patented the silicon chip integrated circuit.
1963	H.B. Maynard: *Industrial Engineering Handbook*.
	Gerald Nadler: *Work Design*.
1964	Abraham Maslow: *Motivation and Personality*.
1965	Transistors were fitted into miniaturized "integrated circuits."

(*continued*)

Table 1.1 (Continued) Major Chronological Events Marking the Origin of Industrial and Systems Engineering

Year	Major publications and events
1966	Frederick Hertzberg: *Work and the Nature of Man.*
1968	Roethlisberger: *Man in Organization.*
	US Department of Defense: *Principles and Applications of Value Engineering.*
1969	Shigeo Shingo developed single-minute exchange of dies (SMED).
	Shigeo Shingo introduced preautomation.
	Wickham Skinner: "Manufacturing—Missing link in corporate strategy" article in *Harvard Business Review*.
1971	Taiichi Ohno completed the Toyota production system.
1971	Intel Corporation developed the microprocessor chip.
1973	First annual Systems Engineering Conference of AIIE.
1975	Shigeo Shingo extolled NSP-SS (non-stock production) system.
	Joseph Orlicky: *MRP: Material Requirements Planning*.
1976	IBM marketed the first personal computer.
1980	Matsushita Electric used Mikuni method for washing machine production.
	Shigeo Shingo: *Study of the Toyota Production System from an Industrial Engineering Viewpoint.*
1981	Oliver Wight: *Manufacturing Resource Planning: MRP II*.
1982	Gavriel Salvendy: *Handbook of Industrial Engineering.*
1984	Shigeo Shingo: *A Revolution in Manufacturing: The SMED System*.

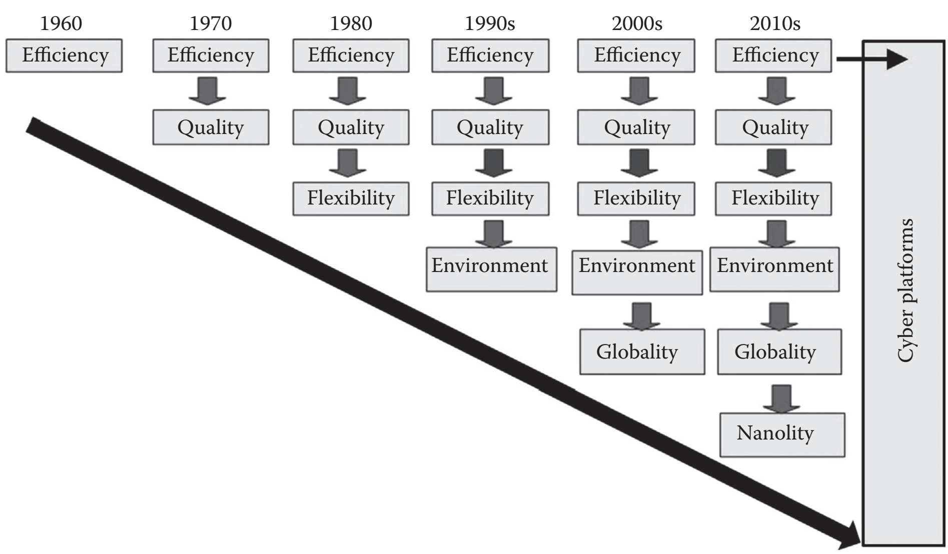

Figure 1.1 Industry progress from classical efficiency to cyber operations.

Table 1.2 Classical Scientific Management Compared with Contemporary Techniques

Taylor's classical principles of scientific management	Equivalent contemporary principles, tools, and techniques	Applicability for ISE skills set
Time studies	Work measurement; process design; PDCA; DMAIC	Effective resource allocation; schedule optimization
Functional supervision	Matrix organization structure; SMART task assignments; lean principles	Team structure for efficiency; people interfaces
Standardization of tools and implements	Tool bins; interchangeable parts; modularity of components; ergonomics; lean principles	Optimization of resource utilization
Standardization of work methods	Six Sigma processes; OODA loop; lean principles	Reduction of variability
Separate planning function	Task assignment techniques; Pareto analysis; lean principles	Reduction of waste and redundancy
Management by exception	Failure mode and effect analysis (FMEA); project management; Pareto analysis	Focus on vital few; task prioritization
Use of slide rules and similar time-saving devices	Blueprint templates; computer hardware and software	Use of boilerplate models
Instruction cards for workmen	Standards maps; process mapping; work breakdown structure; lean principles	Reinforcement of learning
Task allocation and large bonus for successful performance	Benefit–cost analysis; value-added systems; performance appraisal	Cost reduction; productivity improvement; consistency of morale
Use of differential rate	Value engineering; work rate analysis; AHP; lean principles	Input–output task coordination
Mnemonic systems for classifying products and implements	Relationship charts group technology; charts and color coding	Goal alignment; work simplification
A routing system	Lean principles; facility layout; PICK chart; D-E-J-I (design, evaluate, justify, integrate)	Minimization of transportation and handling; reduction of procurement cost
A modern costing system	Value engineering; earned value analysis	Cost optimization

Note: AHP, analytic hierarchy process; DEJI, design, evaluate, justify, integrate; DMAIC, define, measure, analyze, improve, control; FMEA, failure mode and effect analysis; OODA, observe, orient, decide, and act; PDCA, plan-do-check-act; PICK, possible, implement, challenge, or kill; SMART, specific, measurable, aligned, realistic, timed.

part two

Fundamentals of industrial engineering

chapter two

Operational efficiency

Chia-Yen Lee and Andrew L. Johnson

Contents

2.1 Introduction

The fields of engineering and management associate *efficiency* with how well a relevant action is performed, that is, "doing things right," and *effectiveness* with selecting the best action, that is, "doing the right thing." Thus, a firm is *effective* if it identifies appropriate strategic goals, and *efficient* if it achieves them with minimal resources. This chapter focuses on *operational efficiency*, or the ability to deliver products and services cost-effectively without sacrificing quality. In this chapter, we investigate a firm's operational efficiency with both queuing models and productivity and efficiency analysis (PEA) methods that identify maximum productivity and measure efficiency as a ratio of observed productivity to maximum productivity. The maximum productivity level serves as a benchmark for desired performance. The methods for analysis will vary depending on the level of analysis. For example, at the micro-level, we measure operational efficiency at points (machine, workstation, laborer) on the shop floor, whereas the macro-level might be at the firm, industry, or national level. We begin by evaluating performance at the operational level, and then applying PEA to aggregate performance at higher levels.

The analysis of productivity and efficiency is associated with production economics, which focuses on assessment and uses an aggregate description of technology to answer questions (Hackman, 2008) such as

- How efficient is the firm in utilizing its inputs to produce its outputs?
- Is the firm using the right mix of inputs or producing the right mix of outputs given prevailing prices?
- How will the firm respond to a price hike in a critical input?
- How efficient is the firm in scaling its operations?
- Has the firm improved its productive capability over time?
- How does the firm compare to its competitors?

Figure 2.1 shows the three levels of production and operational planning and defines the role of PEA. The *strategic level* includes long-term planning issues such as make-or-buy decisions. The *tactical level* describes midterm actions that are done perhaps on a weekly or monthly basis, while the *operational level* emphasizes daily scheduling and shop floor control. PEA supports tactical-level decisions and is part of midterm production planning. PEA provides performance benchmarking and production guidance. It can also provide *ex post* analysis to quantify efficiency for complex production processes that use multiple inputs to generate multiple outputs, or *ex ante* analysis to suggest guidelines for resource allocation.

2.1.1 *Absolute operational efficiency*

Ideal benchmarks to measure efficiency are usually developed in a design laboratory under perfect operating conditions. However, it is not easy to identify the sources of efficiency loss between ideal performance and the best observed performance. For instance, in a manufacturing process operating in perfect conditions, one machine's ideal throughput is 100 units per hour, yet the actual throughput is 80 units per hour due to operator's skill, scheduling, etc. We can estimate an absolute operational efficiency (AOE) as

$$\text{AOE} = \frac{\text{Actual throughput}}{\text{Ideal throughput}} = \frac{80}{100} = 0.8$$

Note that ideal benchmarks can be observed at the machine or process level, but are almost never observed at the firm level. Thus, alternative metrics are beneficial in the cases when ideal benchmarks are not observable.

2.1.2 *Relative operational efficiency*

Relative operational efficiency (ROE) is the ratio of actual throughput to the best observed throughput. Relative benchmarks are often used to measure efficiency because similar comparable machine, process, firm, etc., are often easily identifiable. We estimate ROE by identifying the best observed performance in a data set of multiple operations performing the same task, for instance, a data set of multiple machines performing the same manufacturing process. We find that the best observed throughput is 90 units per hour, but machine A produces 80 units per hour. We can estimate the ROE of machine A as

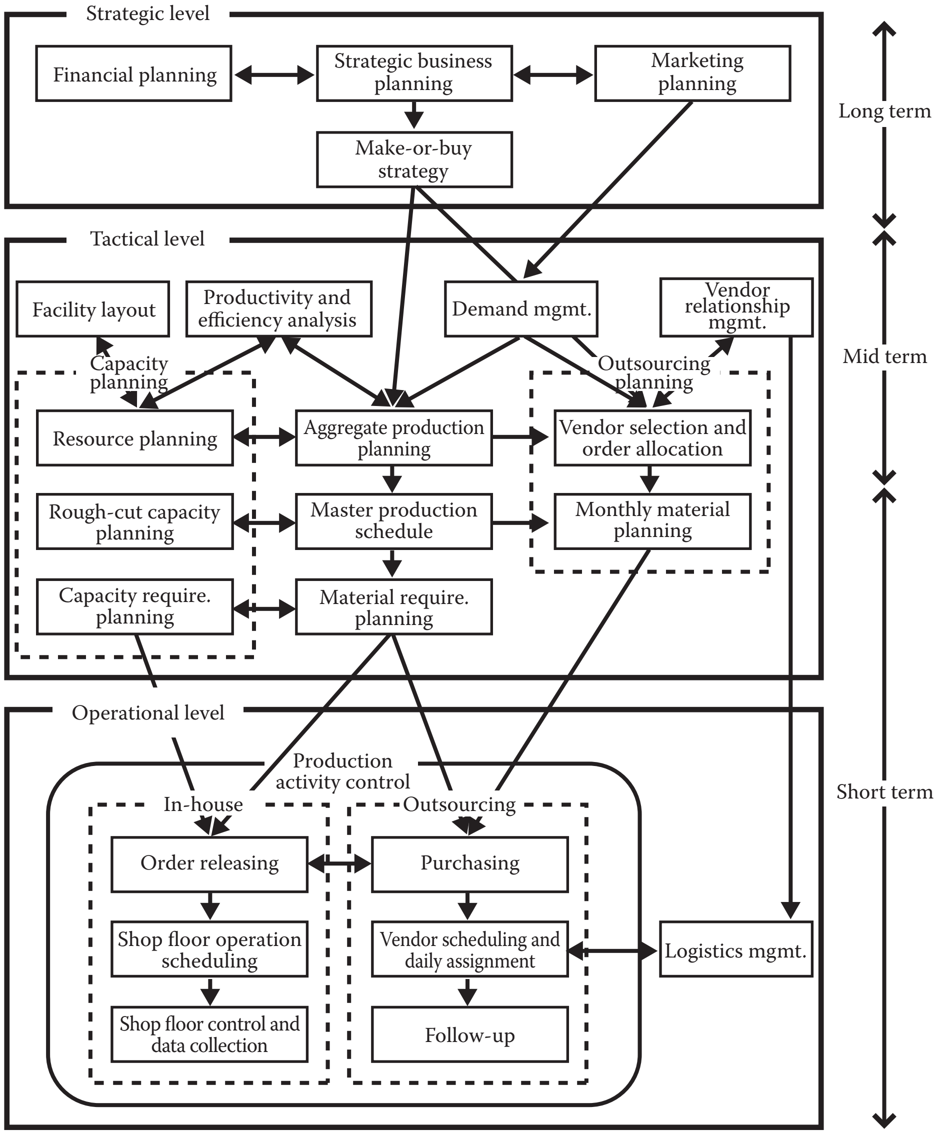

Figure 2.1 General description of analysis levels in production and operations planning.

$$\text{ROE} = \frac{\text{Actual throughput}}{\text{Best observed throughput}} = \frac{80}{90} = 0.88$$

The best observed throughput is often determined by using historical performance data under the assumption, if all conditions are unchanged, that actual throughput should be equal to/or close to the historically best performance.

In the real world, a firm's resources are always limited. When a firm would like to provide a product or service, it must consume input resources to generate the output level. In this setting, operational efficiency is determined by the outputs produced as well as the input resources or costs consumed. Thus, we can define productivity and efficiency as

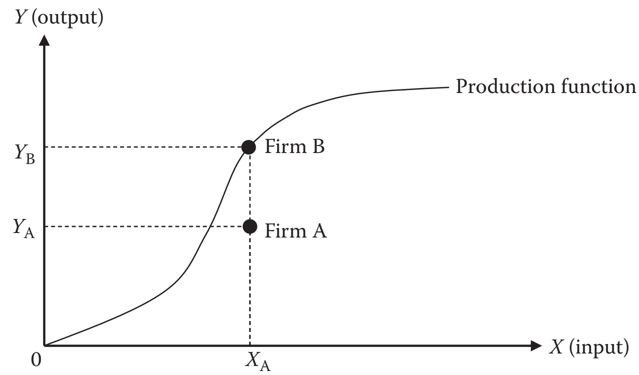

Figure 2.2 Production function and efficiency estimation.

$$\text{Productivity} = \frac{\text{Output}}{\text{Input}}$$

$$\text{Efficiency} = \frac{\text{Productivity}}{\text{Productivity of best practice}}$$

In other words, productivity is the ratio of output level to the input level and efficiency is the ratio of the current productivity level to the best practice productivity level. Best practice is defined as the largest productivity achievable.

The relationship between the output levels produced as input levels change is the production function. Figure 2.2 shows an S-shaped production function with a single input and a single output. We say that firm A is technically inefficient because, given the same input level, firm B is able to produce more output than firm A. We can also say that firm B is efficient because, holding the input level fixed, it produces the highest possible output level. The concept of production function is explained further in Section 2.2.2.

2.2 Efficiency evaluation and performance indices

This section describes efficiency evaluation and related performance indices. Section 2.2.1 discusses how to evaluate efficiency by the queuing theory in the shop floor level. Section 2.2.2 discusses the use of a production function characterization of aggregate performance at the system or firm level as the production process becomes larger and includes workers with uncertain behavior and longer time horizons. Section 2.2.3 introduces three approaches, stochastic frontier analysis (SFA), data envelopment analysis (DEA), and stochastic semi-nonparametric envelopment of data (StoNED), to estimate technical (operational) efficiency by using the observed inputs and outputs levels of a set of firms to estimate a production function.

2.2.1 Shop floor performance and queuing theory

At the shop floor level, queuing models provide a method for evaluating machine performance. In the model shown in Figure 2.3, we use the notation M/M/1 to describe the inter-arrival process, the service process for a single-server queuing system. The first M

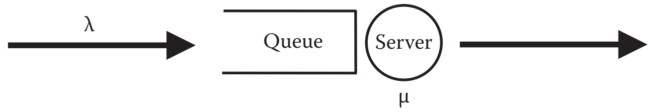

Figure 2.3 M/M/1 queue.

indicates that customer arrivals follow a Poisson (Markovian) Process and the inter-arrival time is exponential distribution. The second M indicates that the service time follows an exponential distribution. The "1" indicates there is a single server.

We use two parameters to describe the M/M/1 queuing system. Let λ be the arrival rate and μ be the service rate. For example, if $\lambda = 2.5$ customers per hour, it means on average 2.5 individuals arrive every hour. Thus, $1/\lambda$ is the mean inter-arrival time and $1/\mu$ is the mean service time. Figure 2.4 shows the Markov state-transition diagram.

The condition $\lambda < \mu$ is necessary for the system to be stable, that is, for the queue to be finite in length. $\rho = \lambda/\mu$ is the probability the server is busy and p_0 is the probability the server is idle. p_i is the probability of the server with i customers. We use the following set of algebraic equations to analyze the queue's performance.

In the beginning, we want to know the stable probability p_0. To characterize a queuing system with the state transition between 0 and 1, a rate-balance equation between the arrival rate and service rate can be shown as

$$\lambda p_0 = \mu p_1 \rightarrow p_1 = \left(\frac{\lambda}{\mu}\right) p_0 = \rho p_0.$$

Intuitively, an empty system needs one arrival to become state 1; a system with one customer needs one departure to become state 0. This idea is the foundation of the rate-balance equation.

Similarly, we can derive the rate-balance equation for state 1 associated with state 0 and state 2.

$$(\lambda + \mu)\, p_1 = \lambda p_0 + \mu p_2 \rightarrow p_2 = (1 + \rho)\, p_1 - \rho p_0 = \rho^2 p_0$$

We can also derive a general formula, $p_n = \rho^n p_0$, for the probability that there are n customers in the system (p_n).

We obtain p_0 since the sum of all probability p_n for $n = 1, \ldots, \infty$ must be equal to 1:

$$\sum_{n=0}^{\infty} p_n = p_0 \sum_{n=1}^{\infty} \rho^n = \frac{p_0}{1-\rho} = 1 \rightarrow p_0 = 1 - \rho$$

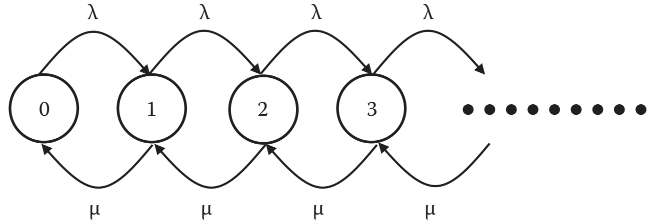

Figure 2.4 Markov-state transition diagram of M/M/1 queuing system.

Thus, we derive the steady-state probability

$$P[\text{server idle}] = p_0 = 1 - \rho$$

$P[\text{server busy}] = 1 - p_0 = \rho = \lambda/\mu$ (also called the "utilization")

$$P[n \text{ customers in the system}] = p_n = \rho^n(1 - \rho)$$

To derive the probability of n or more customers in the system

$$\sum_{m=n}^{\infty} p_m = (1-\rho)\sum_{m=n}^{\infty}\rho^m = (1-\rho)\sum_{k=0}^{\infty}\rho^{n+k}$$
$$= \rho^n(1-n)\sum_{k=0}^{\infty}\rho^k = \rho^n(1-\rho)\frac{1}{1-\rho} = \rho^n$$

$$P[n \text{ or more customers in the system}] = \rho^n$$

$$P[\text{less than } n \text{ customers in the system}] = 1 - \rho^n$$

Thus far, the probability distribution of steady state is derived for a single-server queuing system. We can construct two indices to evaluate the queuing system's performance by asking

What is the expected number of customers in the system/in the queue?
What is the expected time of a customer staying in the system/in the queue?

Let L be the expected number of customers in the system.

$$L = \sum_{n=0}^{\infty} np_n = \sum_{n=0}^{\infty} n\rho^n(1-\rho) = \sum_{n=0}^{\infty} n(\rho^n - \rho^{n+1}) = 1(\rho^1 - \rho^2)$$
$$+2(\rho^2-\rho^3)+3(\rho^3-\rho^4)+\ldots = \rho+\rho^2+\rho^3+\rho^4+\ldots$$
$$= \rho(1+\rho+\rho^2+\rho^3+\ldots) = \frac{\rho}{1-\rho} = \frac{\lambda}{\mu-\lambda}$$

The expected number of customers in the queue, L_q, can be derived similarly. Note that we assume the customer being served is not in the queue, so n customers in the system means the queue length is $n - 1$.

$$L_q = \sum_{n=1}^{\infty}(n-1)p_n = \sum_{n=1}^{\infty} np_n - \sum_{n=1}^{\infty} p_n = L - (1-p_0) = \frac{\rho}{1-\rho}$$
$$-(1-(1-\rho)) = \frac{\rho}{1-\rho} - \rho$$

$$L_q = L - \rho = L\rho$$

Let W be the expected time spent in the system by a customer. Intuitively, it is equal to the expected number of customers in the system divided by arrival rate λ. The equation is

$$L = \lambda W$$

This equation, or Little's Law, defines the relationship between L and W. Similarly, $L_q = \lambda W_q$, where W_q denotes the expected time spent in the queue by a customer

$$W = \frac{L}{\lambda} = \frac{L\rho}{\lambda\rho} = \frac{L_q/\lambda}{\rho} = \frac{W_q}{\rho} \rightarrow W_q = \rho W$$

$$W = \frac{L}{\lambda} = \frac{L_q + \rho}{\lambda} = \frac{\mu L_q + \lambda}{\lambda\mu} = \frac{L_q}{\lambda} + \frac{1}{\mu} = W_q + \frac{1}{\mu}$$

The relationship between W and W_q results because the expected time spent in the system is equal to the expected time spent in the queue plus the mean service time.

Above, Little's Law is defined for a general queuing system. In a manufacturing system, Little's Law is interpreted as the relationship among work-in-process (WIP), throughput (TH), and cycle time (CT)

$$\text{WIP} = \text{TH} \times \text{CT}$$

WIP is the number of unfinished units in the production system, TH is the number of finished products manufactured per unit of time, and CT is the amount of time the units remain in the production system. Given a fixed WIP, an inverse relationship characterizes TH and CT, that is, an increase in TH will decrease CT.

Little's Law is useful because it applies to a wide variety of production systems. Given a fixed TH, WIP and CT will maintain an almost linear relationship until the capacity limit is approached; however, if WIP continues to increase, CT will deteriorate rapidly. Figure 2.5, an example of a workstation, shows that when utilization approaches 100%, the

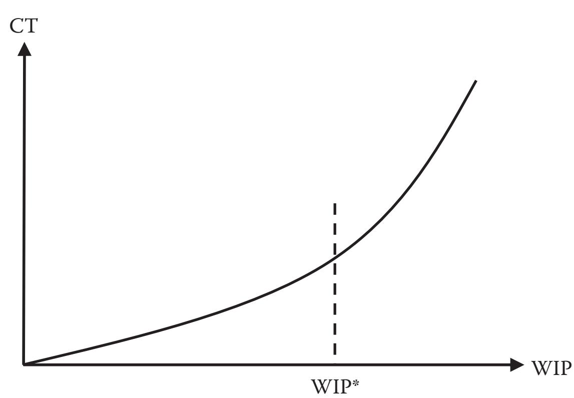

Figure 2.5 CT deterioration

increase of arrival rate λ will deteriorate WIP or CT. Thus, $\lambda > \mu$ implies the workstation is no longer stable.

The typical performance metrics for queuing systems are utilization and throughput. We calculate utilization as

$$\text{Utilization} = \frac{\lambda}{\mu} = \frac{\text{Actual throughput}}{\text{Theoretical (ideal) throughput}}$$

Given CT and the level of WIP, we use Little's Law to calculate the M/M/1 system's productivity by dividing TH by 1. More complicated network analyses are possible with multiple processors linked in a network (Gautam, 2012). Queuing theory can be used to calculate throughput, and productivity can be estimated by dividing throughput by the number of processors. However, all processors may not be identical and throughput will clearly be affected by the underlying network structure. Furthermore, the human component of operating machines adds additional complications and uncertainty that are difficult to capture in queuing models. Thus production functions are useful for estimating complex systems or firm level performance.

2.2.2 Production function

A production function $f(\mathbf{x})$ is the maximum outputs that can be achieved using input vector $\mathbf{x} = (x_1, \ldots, x_N)$ (Hackman, 2008). Outputs are units a firm generates and inputs are the factors of production, or the commodities used in production. In economics, there are at least five types of factors of production: capital, labor, land, energy, and raw materials. We can analyze the performance of a firm's production system in using either the long-run production function or the short-run production function. In the short run, the factors can be divided into fixed factors and variable factors. Fixed factors are the factors that cannot be changed in the short run, such as building and land, and variable factors are the factors that can be changed in the short run, such as temporary workers. In the long run, all of the production factors are variable.

Theoretically, four properties characterize a production function (Chambers, 1988; Coelli et al., 2005):

Non-negativity: The production output is a finite, non-negative, real number.
Weak essentiality: The production output cannot be generated without the use of at least one input.
Monotonicity: Additional units of an input will not decrease output; also called non-decreasing in $\mathbf{x}$.
Concavity: Any linear combination of the vectors $\mathbf{x}^0$ and $\mathbf{x}^1$ will produce an output that is no less than the same linear combination of $f(\mathbf{x}^0)$ and $f(\mathbf{x}^1)$. That is, $f(\lambda\mathbf{x}^0 + (1-\lambda)\mathbf{x}^1) \geq \lambda f(\mathbf{x}^0) + (1-\lambda)f(\mathbf{x}^1)$. This property implies the "law of diminishing marginal returns."

These properties can be relaxed to model-specific production behaviors. For example, monotonicity is relaxed to model input congestion (Färe et al., 1985, 1994)* and concavity is relaxed to characterize an S-shaped production function (Frisch, 1964; Henderson and Quandt, 1980).

* Input congestion indicates that the output level may decrease even though we increase more input due to a difficulty of management and organization.

2.2.2.1 Short-run production function

Because of the fixed factors in the short run, the production function is characterized by monotonically increasing levels and diminishing returns; that is, increasing one variable factor of production will increase output levels at a decreasing rate while holding all others constant. The fixed factors limit the growth of the output. This is also called the law of diminishing marginal returns (product).

Three concepts of production characterize a short-run production function:

Total product (TP): the total amount of output generated from the production system, $\mathrm{TP} = y = f(\mathbf{x})$.

Average product (AP): the average amount of output per unit input, $\mathrm{AP} = \dfrac{f(\mathbf{x})}{\mathbf{x}}$.

Marginal product (MP): the marginal change while adding one more unit of input, $\mathrm{MP} = \dfrac{\mathrm{d}f(\mathbf{x})}{\mathrm{d}\mathbf{x}}$.

Figure 2.6 illustrates a single-input and single-output production function when all other factors are fixed. As the firm increases its input levels, the output levels also increase. The firm

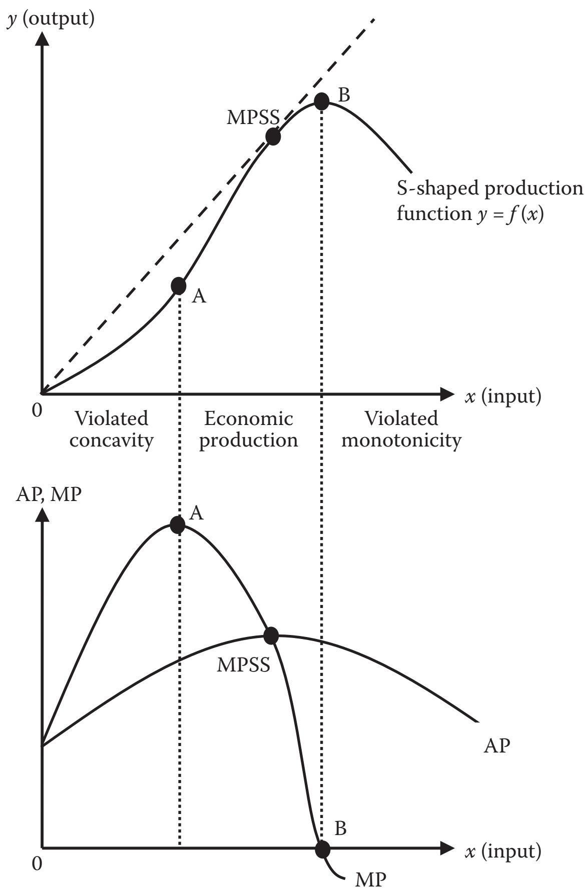

Figure 2.6 Single-input and single-output production function.

reaches point A, an inflection point, that is, where the maximal marginal product is achieved. As inputs continue to increase, the single-input and single-output production function shows diminishing marginal product as it reaches the most productive scale size (MPSS). MPSS is the point on the production function that maximizes the average product (or productivity). Finally, input and output levels continue to increase until point B, beyond which input congestion occurs due to the fixed factors and results in a negative marginal product.

2.2.2.2 Long-run production function

All of the factors of production are variable in the long run. Consider production using multiple inputs. It is common practice to plot the relationship between two of the variables while holding all others fixed. Figure 2.7 shows the relationship between the inputs x_n and x_m while holding the output fixed at the value y^0 and holding all other inputs fixed. The resulting curve is the input isoquant, which gives all combinations of x_n and x_m capable of producing the same output level y^0. It is convex toward the origin if it satisfies all properties of the production function. For different output levels $y^2 > y^1 > y^0$, these isoquants form non-intersecting functions. The slopes of the isoquants are the marginal rate of technical substitution (MRTS), which measures the rate of using x_n to substitute x_m while holding the output level constant:

$$\mathrm{MRTS}_{nm} = -\frac{\partial x_m(x_1,\ldots,x_{m-1},x_{m+1},\ldots,x_N)}{\partial x_n} = \frac{\mathrm{MP}_n}{\mathrm{MP}_m}$$

$$\mathrm{MP}_n\,\partial x_n + \mathrm{MP}_m \partial x_m = 0$$

where $x_m(x_1, \ldots, x_{m-1}, x_{m+1}, \ldots, x_N)$ is an implicit function indicating how much x_m is needed to produce the same output level given fixed levels of $x_1, \ldots, x_{m-1}, x_{m+1}, \ldots, x_N$. Thus, the rate of substitution of input m for input n along the isoquant is equal to the ratio of the marginal productivity of n relative to the marginal productivity of m. To remove the unit of measurement, the direct elasticity of substitution (DES) is the percentage change in the input ratio relative to the percentage change in the MRTS, and quantifies the curvature of the isoquant.

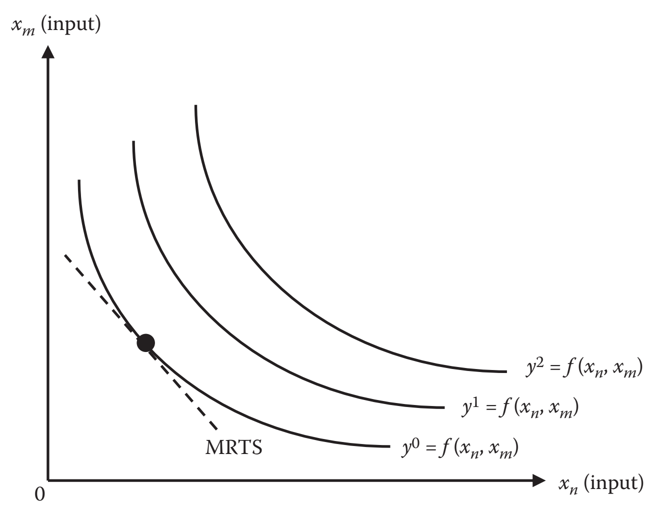

Figure 2.7 Input isoquants.

$$\mathrm{DES}_{nm} = \frac{d(x_m/x_n)}{d(\mathrm{MP}_n/\mathrm{MP}_m)} \times \frac{\mathrm{MP}_n/\mathrm{MP}_m}{x_m/x_n}$$

2.2.2.3 *Three typical production functions for a two-input case*

2.2.2.3.1 Leontief production function. Leontief production functions or fixed-proportions functions describe production that occurs in fixed proportions, for example, cars that require wheels (x_n) and bodies (x_m). The mathematical form is y = min{$\beta_n x_n$, $\beta_m x_m$} and β_n, $\beta_m > 0$; Figure 2.8a shows how the horizontal part of the isoquant indicates that an increase in x_n does not contribute to the output (y), and $\mathrm{MP}_n = 0$ and $\mathrm{MRTS}_{nm} = 0$, and that the vertical part of the isoquant indicates that an increase in x_m does not contribute to the output (y), and $\mathrm{MP}_m = 0$ and $\mathrm{MRTS}_{nm} = \infty$. MRTS_{nm} is not defined at the corner. Therefore, a Leontief production function is used to model production where there is no substitution between x_n and x_m, that is, $\mathrm{DES}_{nm} = 0$.

2.2.2.3.2 Linear production function. A linear production function assumes that inputs are substituted at a constant rate regardless of the level of either input or output. The mathematical form is $y = \beta_n x_n + \beta_m x_m$ and β_n, $\beta_m > 0$; Figure 2.8b shows that the production function implies a constant rate of substitution, $\mathrm{MRTS}_{nm} = \frac{\mathrm{MP}_n}{\mathrm{MP}_m} = \frac{\beta_n}{\beta_m}$, and also imposes perfect substitution between x_n and x_m, that is, $\mathrm{DES}_{nm} = \infty$.

2.2.2.3.3 Cobb–Douglas production function. A Cobb–Douglas production function assumes that inputs are substitutable. However, consistent with the law of diminishing marginal productivity, additional inputs are needed to maintain the same output level as the mix of inputs becomes more skewed. The mathematical form is $y = \alpha x_n^{\beta_n} x_m^{\beta_m}$ and α, β_n, $\beta_m > 0$; Figure 2.7 shows that the production function is a smooth curve and convex toward the origin, and that $\mathrm{MRTS}_{nm} = \frac{\mathrm{MP}_n}{\mathrm{MP}_m} = \frac{\alpha\beta_n x_n^{\beta_n - 1} x_m^{\beta_m}}{\alpha\beta_m x_n^{\beta_n} x_m^{\beta_m - 1}} = \frac{\beta_n x_m}{\beta_m x_n}$ decreases with respect to x_n. Thus, substitution exists in this production function and $0 < \mathrm{DES}_{nm} < \infty$.

2.2.2.4 *Properties of production function*

Figure 2.8 shows that the production functions are convex toward the origin because the absolute value of the slope of the isoquant decreases while increasing x_n; thus, MRTS_{nm}

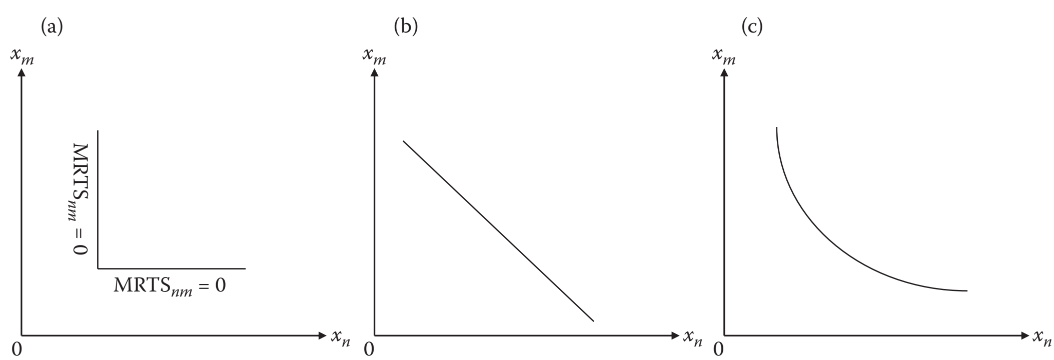

Figure 2.8 Production function for a two-input case.

Table 2.1 Returns to Scale

Return to scale	Mathematical formulation
Decreasing returns to scale (DRS)	$f(\lambda \mathbf{x}) < \lambda f(\mathbf{x})$
Constant returns to scale (CRS)	$f(\lambda \mathbf{x}) = \lambda f(\mathbf{x})$
Increasing returns to scale (IRS)	$f(\lambda \mathbf{x}) > \lambda f(\mathbf{x})$

also decreases. This is called the law of diminishing marginal rate of technical substitution. The mathematical representation is $\frac{\partial}{\partial x_n}\mathrm{MRTS}_{nm} < 0$.

In addition, if a proportionate increase in all inputs results in a less than proportionate increase in output, we say that the production function exhibits decreasing returns to scale (DRS). Alternatively, if increasing all inputs results in the same proportional increase in output, we say that it exhibits constant returns to scale (CRS). Finally, if the increase of all inputs results in a more than proportionate increase in output, we say that the production function exhibits increasing returns to scale (IRS). Table 2.1 shows a mathematical illustration of these three properties where $\lambda > 1$.

There are many reasons why firms may exhibit different returns to scale. For example, a firm may exhibit IRS if hiring more personnel allows specialization of labor; however, the firm may eventually exhibit DRS if the firm becomes so large that management is no longer able to control operations. Firms that can replicate all aspects of their operations exhibit CRS. Operating at decreasing returns to scale would indicate decentralization or downsizing might be appropriate, whereas operating at increasing returns to scale would indicate mergers, acquisitions, or other changes in organizational structure might be appropriate.

2.2.3 Firm-level performance and efficiency estimation

We construct the production function to define a benchmark to measure how efficiently production processes use inputs to generate outputs. Given the same level of input resources, inefficiency is indicated by lower levels of output. In a competitive market, if a firm is far from the production function and operates inefficiently, it needs to increase its productivity to avoid going out of business.

Production theory provides a useful framework to estimate the production function and efficiency levels of a firm in three ways: (1) using parametric functional forms in regression-based methods, for example, SFA (Aigner et al., 1977; Meeusen and van den Broeck, 1977), (2) using nonparametric linear programming methods, for example, DEA (Charnes et al., 1978; Banker et al., 1984), or (3) integrating regression and programming methods, for example, StoNED (Kuosmanen and Kortelainen, 2012; Kuosmanen and Johnson, 2010). In this section, we describe how to use the three methods to estimate efficiency based on cross-sectional data for K firms.

2.2.3.1 Stochastic frontier analysis

Aigner and Chu (1968) use the logarithmic form of the Cobb–Douglas production function to estimate a deterministic frontier

$$\ln y_k = x_k'\beta - u_k$$

where $k = 1, \ldots, K$ and y_k indicates the single output of the firm k; x_k is an $I \times 1$ vector with the elements of logarithm inputs; β is a vector of unknown parameters; and u_k is a non-negative random variable associated with technical inefficiency. Several methods can be used to estimate the parameter β, such as maximum likelihood estimation (MLE) or ordinary least squares (OLS) (Richmond, 1974). However, the Aigner and Chu method neglects statistical noise and assumes that all deviations from the frontier are a result of technical inefficiency. Therefore, Aigner et al. (1977) and Meeusen and van den Broeck (1977) proposed the stochastic frontier production function and introduced the random variable representing statistical noise as

$$\ln y_k = x_k'\beta + v_k - u_k$$

where v_k models the statistical noise using a symmetric random error. The function is bounded from above due to the stochastic variable $\exp(x_k'\beta + v_k)$. To illustrate, we use a Cobb–Douglas stochastic frontier model with a single input variable

$$\ln y_k = \beta_0 + \beta_1 \ln x_k + v_k - u_k$$

$$y_k = \exp(\beta_0 + \beta_1 \ln x_k) \times \exp(v_k) \times \exp(-u_k)$$

In this functional form, $\exp(\beta_0 + \beta_1 \ln x_k)$ is the deterministic component, $\exp(v_k)$ is the statistical noise, and $\exp(-u_k)$ is the inefficiency component. Figure 2.9 illustrates the deterministic frontier $y_k = \exp(\beta_0 + \beta_1 \ln x_k)$, the noise effect, and the inefficiency effect of firm A and firm B. Firm A has a negative random noise component, whereas firm B has a positive noise random noise component. The observed output level is $y_k = \exp(\beta_0 + \beta_1 \ln x_k + v_k - u_k)$ and the frontier output level (i.e., without the inefficiency effect) is $y_k^* = \exp(\beta_0 + \beta_1 \ln x_k)$. The observed output of firm B lies below the deterministic part of the frontier, because the sum of the noise and inefficiency is negative.

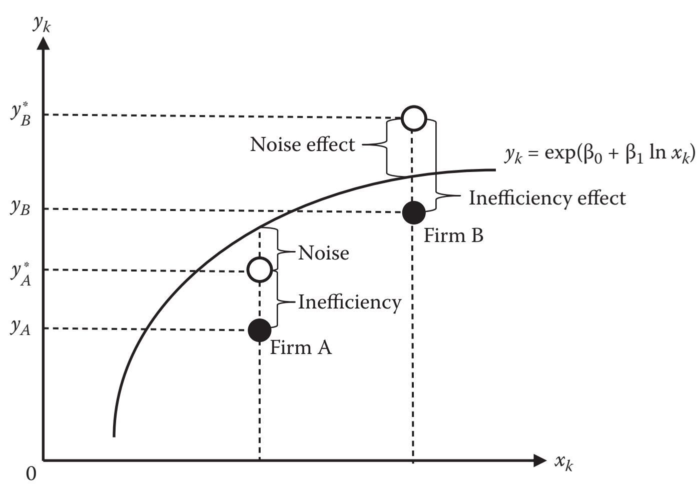

Figure 2.9 Example of stochastic frontier analysis estimate of production function.

We can define the output-oriented measure of technical efficiency (TE) by using the observed output over the frontier output

$$\mathrm{TE}_k = \frac{y_k}{\exp(x_k'\beta + v_k)} = \frac{\exp(x_k'\beta - u_k)}{\exp(x_k'\beta)} = \exp(-u_k)$$

This TE_k estimate shows the measure of observed output of firm k relative to the frontier output of an efficient firm given the same input vector. This benchmarking with best practice provides the estimation of technical inefficiency.

We need to estimate the parameter vector β before calculating TE. Note that the model is complicated by the two random terms, v_i and u_i, where v_i is usually a symmetric error and u_i is a non-negative term. The parameter β is estimated under the following assumptions:

$$E(v_k u_1) = 0,\ \forall k,\ 1\text{: uncorrelated}$$

$$E(v_k) = 0\text{: zero mean}$$

$$E\left(v_k^2\right) = \sigma_v^2\text{: homoskedastic}$$

$$E(v_k v_1) = 0,\ \forall k, \neq 1\text{: uncorrelated}$$

$$E\left(u_k^2\right) = \text{constant: homoskedastic}$$

$$E(u_k u_1) = 0,\ \forall k, \neq 1\text{: uncorrelated}$$

Further, v_k and u_k are uncorrelated with the explanatory variables x_k. Note that $E(u_k) \neq 0$ since $u_k \geq 0$.

To estimate β, Aigner et al. (1977) assume $v_k \sim N\left(0, \sigma_v^2\right)$ and $u_k \sim N^+\left(0, \sigma_u^2\right)$, where v_k follows the independently and identically distributed (iid) normal distribution with zero mean and variance σ_v^2, and u_k follows the iid half-normal distribution, with zero mean and variance σ_u^2. This is called the "half-normal model" in SFA. Under these assumptions, the OLS estimator will provide consistent estimators of slope in β but a downward-biased intercept coefficient since $E(u_k) \neq 0$. Therefore, we use the maximum likelihood estimator (MLE) on the log-likelihood function with $\sigma^2 = \sigma_v^2 + \sigma_u^2$ and $\xi^2 = \sigma_u^2/\sigma_v^2$

$$\ln L(y \mid \beta, \sigma, \lambda) = -\frac{K}{2}\ln\left(\frac{\pi\sigma^2}{2}\right) + \sum_{k=1}^{K} \ln\Phi\left(-\frac{\varepsilon_k \xi}{\sigma}\right) - \frac{1}{2\sigma^2}\sum_{k=1}^{K} \varepsilon_k^2$$

where y is a vector of log-outputs, $v_k - u_k = \ln y_k - x_k'\beta$ defines a composite error term ε_k and Φ is a cumulative distribution function of the stand normal random variable. Finally, we use the iterative optimization procedure to estimate the coefficient β (Judge et al., 1985).

2.2.3.2 Data envelopment analysis

DEA is an optimization-based approach that imposes the axiomatic assumptions of monotonicity and convexity and the minimum extrapolation principle (MEP) (Banker et al., 1984). MEP identifies the smallest set that satisfies the imposed production assumptions and envelops all the data. Thus, DEA estimates a piecewise linear production function based on the observed data points.

Figure 2.10a illustrates 15 production observations and Figure 2.10b illustrates the DEA frontier. The dashed line segment of the DEA frontier represents the strong disposability hull (SDH). That is, the firm on the SDH can decrease the input level without reducing the output level or decrease the output level without changing the input level. We measure the slack in inputs or outputs along the dashed line segments distinguishing the Farrell efficiency measure (Debreu, 1951; Farrell, 1957) and the Koopmans efficiency measure (Koopmans, 1951). The Farrell measure defines technical efficiency as the maximum radial reduction in all inputs consistent with equivalent production of output. The Koopmans measure states that it is impossible for a firm to increase any output without simultaneously reducing another output (or increasing any input). Note that after all inputs have been radially reduced, additional slack may still exist in some but not all inputs. Thus, a Farrell-efficient firm may not be Koopmans efficient.

In this section, we focus on the widely used Farrell measure. First, we introduce the linear programming technique to estimate the production function and production possibility set. Let $x \in R_+^I$ denote the inputs and $y \in R_+^J$ denote the outputs of the production system. We define the production possibility set as $T \equiv \{(x, y)\text{: } x \text{ can produce } y\}$. X_{ik} is the ith input resource, Y_{jk} is the amount of the jth production output, and λ_k is the multiplier for the kth firm. The following model defines the feasible region of the production possibility set $\tilde{T}$. This is called the variable return to scale (VRS) DEA model (Banker et al., 1984) because decreasing marginal product is observed along the frontier

$$\tilde{T} = \left\{(x, y) : \sum_k \lambda_k Y_{jk} \geq Y_j, \forall j; \sum_k \lambda_k X_{ik} \leq X_i, \forall i; \sum_k \lambda_k = 1; \lambda_k \geq 0, \forall k\right\}$$

We use the DEA estimator to measure the efficiency. We describe the input-oriented technical efficiency (ITE) as measured using the distance function $D_x(x, y) = \inf\{\theta \mid (\theta x, y) \in \tilde{T}\}$.

Input-oriented DEA efficiency model

$$\min_\theta \left\{\theta \mid \sum_k \lambda_k Y_{jk} \geq Y_j, \forall j; \sum_k \lambda_k X_{ik} \leq \theta X_i, \forall i; \sum_k \lambda_k = 1; \lambda_k \geq 0, \forall k\right\}$$

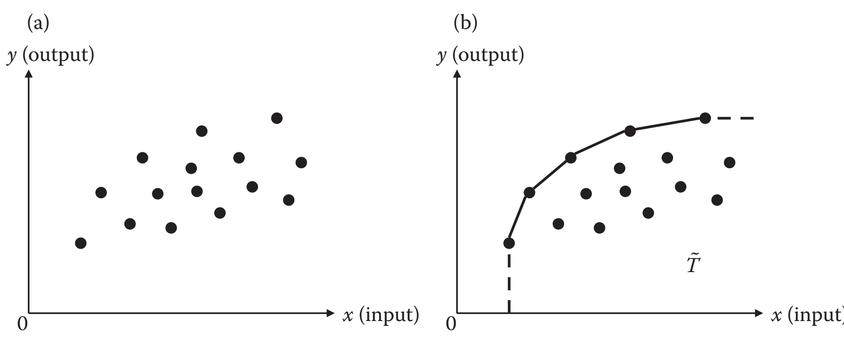

Figure 2.10 DEA frontier with 15 observations.

Output-oriented DEA efficiency model

$$\max_{\omega}\left\{\omega \mid \sum_k \lambda_k Y_{jk} \geq Y_j\omega, \forall j; \sum_k \lambda_k X_{ik} \leq X_i, \forall i; \sum_k \lambda_k = 1; \lambda_k \geq 0, \forall k\right\}$$

We calculate $\theta = 1/\omega$ from the output-oriented DEA efficiency model *i* to get an output-oriented technical efficiency (OTE), θ. $\theta = 1$ implies an efficient firm and $\theta < 1$ implies an inefficient firm.

Figure 2.11 illustrates the input-oriented efficiency measure. Three firms, A, B, and C, are located in an input space constructed by holding the output level constant at $y = \bar{y}$. The solid line is the piecewise linear efficient frontier estimated by DEA. Firms B and C are located on the frontier, but firm A is on the interior of the estimated PPS, $\tilde{T}$. Using the Farrell measure to estimate the technical efficiency shows that the inputs of firm A can be reduced radially. Point D is the intersection of the line segments $\overline{OA}$ and $\overline{BC}$. In fact, point D is a convex combination of firms B and C. We estimate firm A's technical efficiency as

$$TE_A = \theta = D_x\left(x_A, y_A\right) = \frac{\overline{OD}}{\overline{OA}}$$

2.2.3.3 Stochastic semi-nonparametric envelopment of data

The benefits of both SFA and DEA can be achieved using the nonparametric regression approach, StoNED. The first stage of StoNED uses convex nonparametric least squares (CNLS) proposed by Hildreth (1954) and extended by Hanson and Pledger (1976) to estimate a function satisfying continuity, monotonicity, and global concavity—the standard regularity conditions for a production function. To include both random noise and technical inefficiency, Kuosmanen and Kortelainen (2012) combine the CNLS piecewise linear production function with the composite disturbance term concept from SFA.

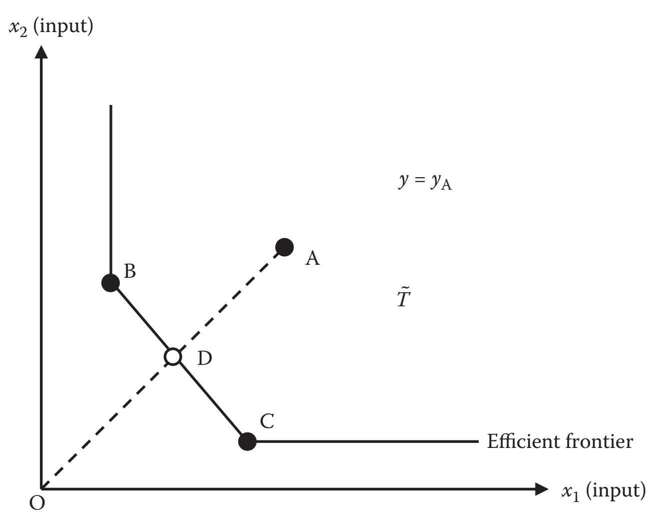

Figure 2.11 Efficiency estimation relative to a DEA input isoquant.

Let $x_k \in R_+^I$ be an input vector, $y_k \in R_+$ be an output, and f be an unknown frontier production function satisfying continuity, monotonicity, and concavity. The regression model is

$$y_k = f(x_k) + \varepsilon_k \quad \forall k = 1, \ldots, K$$

where ε_k is a disturbance term with $E(\varepsilon_k) = 0\ \forall k$, $\mathrm{Var}(\varepsilon_k) = \sigma^2 < \infty\ \forall i$ and $\mathrm{Cov}(\varepsilon_k \varepsilon_j) = 0\ \forall k \neq j$. We formulate the CNLS problem as the quadratic program

$$\min_{\alpha,\beta,\varepsilon} \sum_k \varepsilon_k^2$$

s.t.

$$\varepsilon_k = y_k - (\alpha_k + x_k'\beta_k) \quad \forall k = 1, \ldots, K$$

$$\alpha_k + x_k'\beta \leq \alpha_h + x_k'\beta_h \quad \forall h, k=1, \ldots, K$$

$$\beta_k \geq 0 \quad \forall k = 1, \ldots, K$$

where α_k and β_k are the coefficients characterizing the hyperplanes of the frontier production function f. Note that α_k and β_k are specific to each firm k. The objective function minimizes the sum of squared disturbance terms. The equality constraint defines the disturbance term as the difference between an observed output and an estimated output. The inequality constraints comprise a system of Afriat inequalities (Afriat, 1972), imposing the underlying frontier production function to be continuous and concave. The last constraints enforce monotonicity. Unlike DEA, CNLS uses all of the data points to estimate a production function, making it more robust to outliers.

The CNLS estimator of the production function, $\hat{f}(x)$, is generally not unique, but the fitted output values at observed inputs, $\hat{f}(x_k)$, are unique (Kuosmanen, 2008). In fact, given the fitted output values, it is possible to derive the tightest lower bound of the frontier production function as the explicit lower bound representor function

$$\hat{f}_{\min}(x) = \min_{\alpha,\beta} \left\{ \alpha + x_k'\beta \mid \alpha + x_k'\beta \geq \hat{y}_k \quad \forall k = 1, \ldots, K \right\}$$

where $\hat{y}_k = \hat{f}(x_k)$ is the fitted output value. Since the tightest lower bound $\hat{f}_{\min}$ is a piecewise linear function satisfying continuity, monotonicity, and concavity, we can use it as the unique CNLS estimator of the frontier production function f.

StoNED uses a similar approach to SFA for modeling inefficiency and noise terms. Consider the composite disturbance term

$$\varepsilon_k = v_k - u_k \quad \forall k = 1, \ldots, K$$

where the same properties for v_k and u_k are assumed as in the SFA section.

The composite disturbance term violates the Gauss–Markov property that $E(\varepsilon_k) = E(-u_k) = -\mu < 0$; therefore, we modify the composite disturbance term as

$$y_k = [f(x_k) - \mu] + [\varepsilon_k + \mu] = g(x_k) + \vartheta_k \quad \forall k = 1, \ldots, K$$

where $\vartheta_k = \varepsilon_k + \mu$ is a modified composite disturbance with $E(\vartheta_k) = E(\varepsilon_k + \mu) = 0$ and $g(x_k) = f(x_k) - \mu$ is an average production function. Since g inherits the continuity, monotonicity, and concavity, the CNLS method can find the estimator of the average production function g. We formulate the composite disturbance CNLS problem as

$$\min_{\alpha,\beta,\vartheta} \sum_k \vartheta_k^2$$

$$\text{s.t.} \quad \vartheta_k = y_k - (\alpha_k + x_k'\beta_k) \quad \forall k = 1, \ldots, K$$

$$\alpha_k + x_k'\beta_k \le \alpha_h + x_k'\beta_h \quad \forall k, h = 1, \ldots, K$$

$$\beta_k \ge 0 \quad \forall k = 1, \ldots, K$$

where α_k and β_k are the coefficients that characterize the hyperplanes of the average frontier production function g. Note that the composite disturbance CNLS problem only differs from the CNLS problem in that the sum of squared modified composite disturbances is minimized.

To illustrate the StoNED estimator, 100 observations of a single-input single-output Cobb–Douglas production function are generated, $y = x^{0.6} + v - u$. The observations, x, were randomly sampled from a uniform [1,10] distribution, v was drawn from a normal distribution with standard deviation of 0.5, and u was drawn from a half-normal distribution with standard deviation of 0.7. Figure 2.12 shows the StoNED estimator.

The second stage of StoNED uses the modified composite residuals, $\hat{\vartheta}_k \forall k$, to separate the technical inefficiencies and random noises by applying the method of moments (Aigner et al., 1977; Kuosmanen and Kortelainen, 2012). Assuming that technical inefficiency has a half normal distribution, $u_k \sim \left|N\left(0, \sigma_u^2\right)\right|$, and that random noise has a normal distribution, $v_k \sim N\left(0, \sigma_v^2\right)$, the estimated standard deviation of technical inefficiency and random noise is

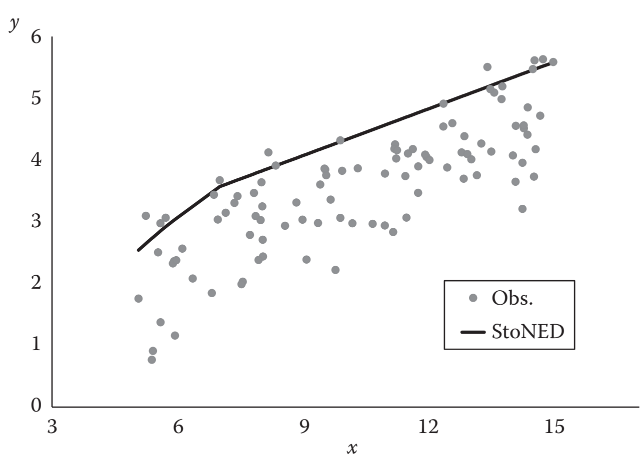

Figure 2.12 StoNED frontier with 100 observations.

$$\hat{\sigma}_u = \sqrt[3]{\frac{\hat{M}_3}{\left(\frac{2}{\pi}\right)\left(1-\frac{4}{\pi}\right)}}$$

$$\hat{\sigma}_v = \sqrt{\hat{M}_2 - \left(\frac{\pi-2}{\pi}\right)\hat{\sigma}_u^2}$$

where $\hat{M}_2 = \frac{1}{n}\sum_k \left(\hat{\vartheta}_k - \hat{E}(\vartheta_k)\right)^2$ and $\hat{M}_3 = \frac{1}{n}\sum_k \left(\hat{\vartheta}_k - \hat{E}(\vartheta_k)\right)^3$ are the second and third sample central moments of the modified composite residuals, respectively. Moreover, $\hat{M}_3$ should be negative so that $\hat{\sigma}_u$ is positive. Intuitively, the composite residuals should have negative skewness reflecting the presence of the technical inefficiency. We calculate the expected technical inefficiency by

$$\hat{\mu} = \hat{\sigma}_u \sqrt{2/\pi} \, .$$

Given $\left(\hat{\alpha}_k, \hat{\beta}_k\right)$ from the CNLS problem, we write the unique StoNED estimator of the frontier production function as

$$\hat{f}_{\min}(x) = \min_{\alpha,\beta} \left\{\alpha + x_k'\beta \mid \alpha + x_k'\beta \geq \hat{y}_k \quad \forall k = 1,\ldots,K\right\} + \hat{\mu}$$

where $\hat{y}_k = \min_{h\in\{1,\ldots,n\}}\left\{\hat{\alpha}_h + x_k'\hat{\beta}_h\right\}$. We obtain the unique CNLS estimator of the average frontier production function, $\hat{g}_{\min}$, by using the tightest lower bound representor function with the fitted output values, $\hat{y}_k$. Recall that $\hat{y}_k$ is calculated from the representor function and $\left(\hat{\alpha}_k, \hat{\beta}_k\right)$. Therefore, we obtain the frontier production function by additively shifting the unique CNLS estimator of the average frontier production function upward by the expected value of technical inefficiency.

Given $\hat{\sigma}_u$ and $\hat{\sigma}_v$, the method introduced in Jondrow et al. (1982) can estimate firm-specific inefficiency. Specifically

$$\hat{E}\left(u_k \mid \hat{\varepsilon}_k\right) = -\frac{\hat{\varepsilon}_k \hat{\sigma}_u^2}{\hat{\sigma}_u^2 + \hat{\sigma}_v^2} + \frac{\hat{\sigma}_u^2 \hat{\sigma}_v^2}{\hat{\sigma}_u^2 + \hat{\sigma}_v^2}\left[\frac{\phi(\hat{\varepsilon}_k/\hat{\sigma}_v^2)}{1-\Phi(\hat{\varepsilon}_k/\hat{\sigma}_v^2)}\right]$$

where $\hat{\varepsilon}_k = \hat{\vartheta}_k - \hat{\mu}$, ϕ is the standard normal density function and Φ is the standard normal cumulative distribution.

2.3 *Efficiency improvement*

Section 2.2 provided models to estimate the system performance and efficiency. This section provides some methodologies for driving productivity. Section 2.3.1 introduces overall equipment effectiveness (OEE) and Section 2.3.2 describes lean manufacturing.

2.3.1 Overall equipment effectiveness

OEE is a time-based metric to assess productivity and efficiency, particularly for the semiconductor manufacturing industry (Ames et al., 1995; Semiconductor Equipment and Material International, 2000, 2001; de Ron and Rooda, 2005). The traditional single index metrics of productivity, throughput, and utilization do not allow easy identification of root cause for reduced productivity. The OEE definition describes six standard equipment states (Figure 2.13):

Non-scheduled state: Equipment is not scheduled to be used in production, such as unworked shifts, weekends, or holidays (including startup and shutdown).

Unscheduled down state: Equipment is not in a condition to perform its intended function owing to unplanned downtime events, for example, maintenance delay, repair, change of consumables or chemicals, and out-of-spec input.

Scheduled down state: Equipment is not available to perform its intended function owing to planned downtime events, for example, production test, preventive maintenance, and setup.

Engineering state: Equipment is in a condition to perform its intended function but is operated to conduct engineering experiments, for example, process engineering, equipment engineering, and software engineering.

Standby state: Equipment is in a condition to perform its intended function but is not operated; the standby state includes no operator available (including breaks, lunches, and meetings), no items available (including no items due to lack of available support equipment), and no support tools.

Productive state: Equipment is performing its intended functions, for example, regular production (including loading and unloading of units), work for third parties, rework, and engineering runs done in conjunction with production units.

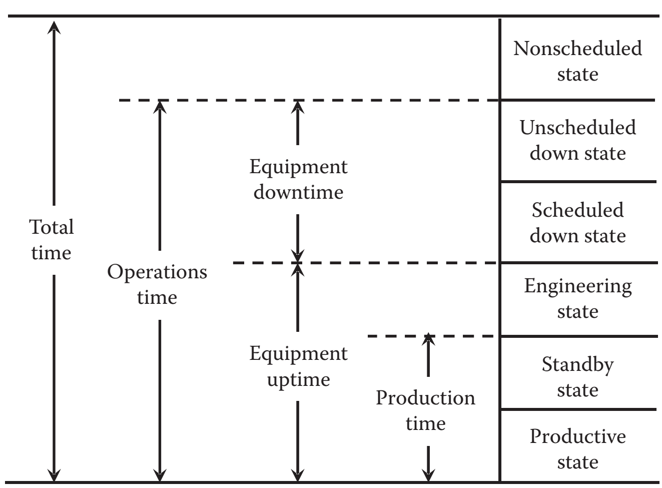

Figure 2.13 OEE and equipment states.

We define OEE as

$$\text{OEE} = \frac{\text{Theoretical production time for effective units}}{\text{Total time}}$$

We decompose OEE into the following subcomponents: availability efficiency (AE), operational efficiency (OE), rate efficiency (RE), and quality efficiency (QE) (de Ron and Rooda, 2005):

$$\text{OEE} = \text{AE} \cdot (\text{OE} \cdot \text{RE}) \cdot \text{QE} = \text{Availability} \cdot \text{Performance} \cdot \text{Quality}$$

where

$$\text{Availability} = \text{AE} = \frac{\text{Equipment uptime}}{\text{Total time}}$$

$$\text{Performance} = \text{OE} \times \text{RE}$$

$$\text{OE} = \frac{\text{Production time}}{\text{Equipment uptime}}$$

$$\text{RE} = \frac{\text{Theoretical production time for actual units}}{\text{Production time}}$$

$$\text{Quality} = \text{QE} = \frac{\text{Theoretical production time for effective units}}{\text{Theoretical production time for actual units}}$$

The availability captures the difference between machine breakdown and processing. Performance characterizes the production time and throughput. The quality is described by the yield metric, which is typically driven by scrap, rework, defects, and reject types. In other words, OEE is a metric to estimate the efficiency of theoretical production time for effective units. In particular, the theoretical production time means the production time without efficiency losses. In addition, two popular indices can be integrated into the OEE framework: mean time between failure (MTBF), or the average time a machine operates before it fails, and mean time to repair (MTTR), or the average time required to repair a failed component and return the machine to operation:

$$\text{AE} = \frac{\text{Equipment uptime}}{\text{Total time}} = \frac{\text{Equipment uptime}}{\text{Equipment uptime} \times \dfrac{(\text{MTTR} + \text{MTBF})}{\text{MTBF}}}$$

$$= \frac{\text{MTBF}}{\text{MTTR} + \text{MTBF}}$$

OEE has two practical benefits. First, we can use its subcomponents to identify bottlenecks and improve productivity. In general, machines with high utilization are typically the bottlenecks. Because bottlenecks can shift depending on the product mix, it is important for engineers to identify and release bottlenecks quickly to maintain high throughput levels. Note that the utilization is a necessary condition for bottleneck identification but does not mean that all high-utilization machines are bottlenecks. If the processing time of each product is the same and the variation in the production line is low, a machine may have high utilization without affecting throughput. Second, we can use OEE to separate a machine's status into regular operating conditions and downtime. The availability level quantifies the time used for production. A lower throughput is sometimes the result of low availability rather than poor performance. Thus, OEE decomposition helps with machine diagnosis and productivity improvement.

2.3.2 Lean thinking and manufacturing

Lean manufacturing has its roots in the manufacturing processes developed by Henry Ford in the 1920s. The Ford Motor Company increased its revenue during the post-World War I depression by developing assembly line methods and eliminating activities that were either unnecessary or did not add value to the cars produced. Toyota coined the name and the concept of lean manufacturing in its production system in the 1980s, and also developed additional supporting methods and concepts such as the just-in-time (JIT) system (Ohno 1988a,b). We call a production system "lean" if it produces the required output levels with minimal buffering costs.

In fact, the only time a machine adds value is when it processes a part. Figure 2.14 provides a Gantt chart to visualize processing time, transportation time, and wait time. Note that loading products into tools is handling, not processing, and thus a non-value-adding activity. Most of the processing time of a product involves waiting and non-value-adding activities. Smith (1998) proposed a manufacturing performance index called manufacturing cycle efficiency:

$$\text{Manufacturing cycle efficiency} = \frac{\text{Value-adding time}}{\text{Total cycle time}}$$

He pointed out that this index is often less than 1% in practice, meaning that firms usually waste resources performing non-value-adding activities.

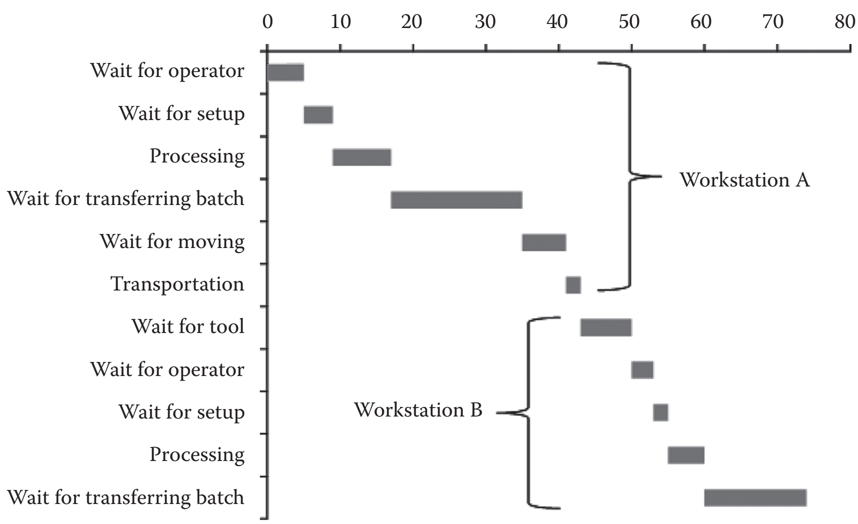

Figure 2.14 Gantt chart of product transition.

The basic philosophy of lean manufacturing is to eliminate the waste by buying only enough material to fit the immediate needs of the production plan considering the transportation time.

Below, we describe the three main principles of lean manufacturing:

Waste elimination
Continuous flow
Pull production system

As the term implies, waste elimination reduces all forms of waste in the manufacturing process. Continuous flow smoothes and balances the production flow. Pull production system, or "make-to-order production," allows a firm to produce units only when it receives an order. There are four steps to implementing lean manufacturing:

1. Eliminate waste: seven types of waste are identified and need to be eliminated.
2. Use buffers: build up, adjust, and swap buffers to manage for variability.
3. Continuous improvement: a commitment to productivity improvement.
4. Reduce variability: identify and reduce internal and external causes.

A firm can allocate resources dynamically and switch buffers to manage internal or external variability. Internal variability results from uncertain processing times, setups, machine breakdown, yield loss, rework, engineering change orders, etc., and external variability results from demand fluctuation, customer change orders, supplier uncertain delivery, etc. Lean manufacturing uses three buffers: inventory, capacity, and time. Inventory hedges against uncertain demand. Capacity is somewhat flexible due to hiring/layoffs of temporary workers, adjusting overtime, or outsourcing some activities. Time coordinates supply chain or manufacturing activities.

The benefits of lean manufacturing include

Productivity improvement
Total manufacturing time saved
Less scrap
Lower inventory
Quality improvement
Plant space saved
Better labor utilization
Lower production cost, higher profits and wages
Shorter cycle time: make-to-order versus make-to-stock
Safety of operations

2.3.2.1 Waste elimination

Womack and Jones (2003) describe seven types of "muda," or waste, in a production system:

Transportation: move products or materials that are not being processed between workstations, or between supplier and customer.
Inventory: hold excess inventory of raw materials, WIP, or finished units.

Motion: worker or tools move more than necessary, such as picking, sorting, stacking, or storing parts.
Waiting: wait for upcoming tools, materials, or parts, or for the next production step.
Overproduction: generate excess products beyond the demand level.
Overprocessing: working more than is necessary because of poor tool or product design.
Defects: cost of poor quality such as rework, scrap, inspection, and repair.

In general, all seven types of waste described above belong to the category of non-value-adding activities. Table 2.2 lists some of the tools Toyota developed to eliminate waste.

2.3.2.2 *Continuous flow*

Continuous flow, or the series of continuous and smooth processes, is the second principle. Each production step performs only the jobs necessary for the next step. Workstations do not hold unnecessary WIP and materials that block incoming and downstream flows. Table 2.3 lists some tools to achieve continuous flow.

2.3.2.2.1 Single-minute exchange of die. Single-minute exchange of die (SMED), or "Shingo," can significantly reduce setup time and improve productivity. Long setup time leads to a small number of setups, larger batch sizes, larger WIP inventories, and poor process flow. SMED divides the setup time into internal and external activities. An internal activity is one that can only be done when the machine is stopped, such as multichamber adjustments; an external activity is anything that can be performed before or after the setup without stopping the machine, such as preheating of raw material. To achieve a

Table 2.2 Tools Developed by Toyota to Eliminate Waste

Tool	Description
Flexible manufacturing	A flexible production system allows quick response to change, in particular, change in product mix. Machine flexibility allows the operator to change the configuration to produce different product types. Routing flexibility allows multiple machines to perform the same function on a product.
Standardize work	Standardize regular operations according to the benchmarking of best practice; post at workstations.
5S	• Seiri (Sort), or "Tidiness": throw away unrelated materials; only leave necessary items at workstation. • Seiton (Set-in-order), or "Orderliness": put everything in its place for quick pick-up and storage. • Seiso (Shine), or "Cleanliness": clean up the workplace. • Seiketsu (Standardize): hold the gains and maintain cleanliness. • Shitsuke (Sustain), or "Discipline": commitment to practice 5S for ongoing improvement.
Automation (Jidoka)	A supervisory function uses automation instruments to detect abnormalities and identify root causes; if an error arises, the production line shuts down immediately; to prevent the defective product and overproduction.
Others	Continuous improvement (Kaizen); error proofing (Poka-yoke); radical change (Kaikaku); worker suggestions (Teien systems); dynamic allocation of workers (Shojinka); etc.

Table 2.3 Tools for Continuous Flow

Tool	Description
Single-minute exchange of die, SMED (Shingo)	Rapid changeover and setup time reduction in converting current manufacturing process to manufacture the next product; improves production flow and reduces lot sizes.
Andon	Uses signboard or visual signals to indicate the location of the alert for abnormality detection.
Takt time	Identifies the allowable time for process steps; calculated by taking available production time over customer demand; used to reduce the gap between current CT and the minimum possible time.
Line balancing	Organize tasks into groups, with each group of tasks being performed at a single workstation; each workstation has identical loading and CT. No workstation is overburdened, no one waits, and the variation is smoothed at each workstation.
Nagara (smooth production flow)	Shortens the lead time between manufacturing processes and reduces WIP inventories to adjust for fluctuations in demand; batch size reduction is a way to reduce inventory and smooth production flow.
Others	Cross-train workers to manage for inherent variability, etc.

quick setup and changeover of dies, SMED recommends reducing internal setup time or converting internal activities to external activities.

2.3.2.2.2 Production line balancing. Line balancing is a typical problem of the assembly system design in industrial engineering (Nof et al., 1997). To compensate for demand fluctuations, the goal is to organize tasks into different groups with each group taking the same amount of time.

The line balancing problem is an NP-hard (non-deterministic polynomial-time hard) problem (Garey and Johnson, 1979); thus, heuristic methods are usually applied to provide good solutions. Helgeson and Birnie (1961) proposed a heuristic method called the ranked positional weight technique. This heuristic is a task-oriented technique considering the combination of precedence relationships and task processing time. Three steps are applied in this algorithm.

1. Calculate the positional weight (PW) of each task using the processing time (PT) of the task plus the processing time of all tasks having this task as a predecessor.
2. Rank tasks in descending order in terms of PW.
3. Assign tasks to workstations sequentially in the ranked order, given the precedence relationships and CT constraint.

Figure 2.15 shows eight tasks with their PT (unit: minute) and the precedence relationships. If the CT is 10 min for each workstation, we calculate the minimal number of workstations according to the sum of the eight task times over the CT, that is, 38/10 = 3.8 and round up to 4. However, this minimum number does not consider the precedence constraints. Thus, we use the ranked PW technique for line balancing as shown in Table 2.4. We find that the required number of workstation is 5 and the total idle time is 12, both of which tend to increase at downstream stations.

The smoothing can be done by product type or by volume; both are quite efficient and can bring substantial efficiencies and savings. Note that a smoothed and continuous flow can be reviewed from a firm's internal production or its supply chain. The benefits include

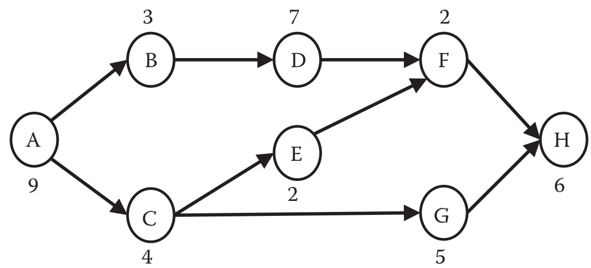

Figure 2.15 Precedence relationships and processing time.

Table 2.4 Ranked Positional Weight Technique

	PT	PW	Order	Station
A	9	38	1	1
B	3	18	3	2
C	4	19	2	2
D	7	15	4	3
E	2	10	6	2
F	2	8	7	3
G	5	11	5	4
H	6	6	8	5

- Enhance flexibility by reducing batch size to accommodate changes in product mix or demand fluctuation.
- Reduce material, WIP, and inventory levels since there is no severe overproduction or underproduction.
- No bottlenecks because of similar burdens for each workstation.
- Enhance loyalty and commitment to the firm, that is, a stable workforce without temporary labor.
- Shorten changeover and setup times to reduce machine idleness.

2.3.2.3 Pull production system

Push systems release work without consideration of system status and hence do not have an inherent limitation on WIP. The work is released on the basis of a schedule of demand and controlled release rates, typically referred to as a due-date-driven production system. A pull system developed by Toyota releases work based on the status of the systems and has an inherent WIP limitation. The system authorizes work releases based on system status and controls WIP level. It is an order-driven production system (Hopp and Spearman, 2004).

There are two techniques in the customer-pull production system: JIT and kanban. JIT attempts to reduce inventory, holding costs, and WIP using a small lot size or even single unit processing. A "kanban" is a signboard for realizing JIT and often leads to significant quality improvement. The advantages of using pull production system include

- Reduce WIP and CT: limit releases into the production line.
- Improve quality: short queues allow errors to be identified quickly and shut down the production line to correct the problems.

- Reduce cost: switch the control from release rate to WIP level and reduce WIP progressively.
- Logistical benefits: less congestion, easier control, and WIP cap control.

Kanban provides for efficient lot tracking and predetermines the WIP level by the number of kanban. In fact, on the basis of Little's Law, WIP = CT × TH, given the same rate of throughput, reducing the WIP level will lead to a reduction in CT. Thus, a pull production system reduces CT by controlling the WIP level. For further study of the pull system, see Ohno (1998a,b), Liker (2004), and Nahmias (2009).

2.4 Conclusion

Operational efficiency can be measured and improved using the approaches described in this chapter. Today, many manufacturing firms define a metric for efficiency and concentrate on operational improvement activities to increase it. The specific approaches developed to identify best practice performance or to determine if a particular activity adds value are often product or industry specific. However, the evolution of new—and global—industries will require more sophisticated efficiency analysis techniques and metrics.

References

Afriat, S.N., Efficiency estimation of production functions, *Int. Econ. Rev.*, 13 (3), 568–598, 1972.

Aigner, D.J. and Chu, S.F., On estimating the industry production function, *Am. Econ. Rev.*, 58, 826–839, 1968.

Aigner, D., Lovell, C.A.K., and Schmidt, P., Formulation and estimation of stochastic frontier production function models, *J. Econom.*, 6, 21–37, 1977.

Ames, V.A., Gililland, J., Konopka, A., Schnabl, R., and Barber, K., Semiconductor manufacturing productivity: Overall equipment effectiveness (OEE) guidelines, Technology transfer 950327443 A-GEN, Revision 1.0. Sematech, 1995.

Banker, R.D., Charnes, A., and Cooper, W.W., Some models for estimating technical and scale inefficiencies in data envelopment analysis, *Manage. Sci.*, 30 (9), 1078–1092, 1984.

Chambers, R.G., *Applied Production Analysis: A Dual Approach*, Cambridge University Press, New York, 1988.

Charnes, A., Cooper, W.W., and Rhodes, E., Measuring the efficiency of decision making units, *Eur. J. Oper. Res.*, 2 (6), 429–444, 1978.

Coelli, T.J., Prasada Rao, D.S., O'Donnell, C.J., and Battese, G.E., *An Introduction to Efficiency and Productivity Analysis*, 2nd ed., Springer, New York, 2005.

Debreu, G., The coefficient of resource utilization, *Econometrica*, 19, 273–292, 1951.

de Ron, A.J. and Rooda, J.E., Equipment effectiveness: OEE revisited, *IEEE Trans. Semicond. Manuf.*, 18 (1), 190–196, 2005.

Färe, R.S., Grosskopf, S., and Lovell, C.A.K., Technical efficiency of Philippine agriculture, *Appl. Econ.*, 17, 205–214, 1985.

Färe, R.S., Grosskopf, S., and Lovell, C.A.K., *Production Frontiers*, Cambridge University Press, Cambridge, 1994.

Farrell, M.J., The measurement of productive efficiency, *J. R. Stat. Soc. A*, 120 (3), 253–281, 1957.

Frisch, R., *Theory of Production*, Rand McNally & Company, Chicago, 1964.

Garey, M.R. and Johnson, D.S., *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman, New York, 1979.

Gautam, N., *Analysis of Queues: Methods and Applications*, CRC Press (Taylor & Francis), Boca Raton, FL, 2012.

Hackman, S.T., *Production Economics: Integrating the Microeconomic and Engineering Perspectives*, Springer-Verlag, Heidelberg, 2008.

Hanson, D.L. and Pledger, G., Consistency in concave regression, *Ann. Stat.*, 4 (6), 1038–1050, 1976.
Helgeson, W.P. and Birnie, D.P., Assembly line balancing using the ranked positional weight technique, *J. Ind. Eng.*, 12, 394–398, 1961.
Henderson, J.M. and Quandt, R.E., *Microeconomic Theory: A Mathematical Approach*, 3rd ed., McGraw-Hill, New York, 1980.
Hildreth, C., Point estimates of ordinates of concave functions, *J. Am. Stat. Assoc.*, 49 (267), 598–619, 1954.
Hopp, W.J. and Spearman, M.L., To pull or not to pull: What is the question? *Manuf. Serv. Oper. Manage.*, 6 (2), 133–148, 2004.
Jondrow, J., Lovell, C.A.K., Materov, I.S., and Schmidt, P., On the estimation of technical inefficiency in the stochastic frontier production function model, *J. Econom.*, 19 (2–3), 233–238, 1982.
Judge, G.G., Griffiths, W.E., Hill, R.C., Lutkepohl, H., and Lee, T.-C., *Introduction to the Theory and Practice of Econometrics*, John Wiley & Sons Inc., New York, 1985.
Koopmans, T., An analysis of production as an efficient combination of activities, in *Activity Analysis of Production and Allocation*, Cowles Commission for Research in Economics, Monograph No. 13, Koopmans, T.C., Ed., John Wiley & Sons Inc., New York, 1951.
Kuosmanen, T., Representation theorem for convex nonparametric least squares, *Econom. J.*, 11, 308–325, 2008.
Kuosmanen, T. and Johnson, A.L., Data envelopment analysis as nonparametric least squares regression, *Oper. Res.*, 58 (1), 149–160, 2010.
Kuosmanen, T. and Kortelainen, M., Stochastic non-smooth envelopment of data: Semi-parametric frontier estimation subject to shape constraints, *J. Prod. Anal.*, 38 (1), 11–28, 2012.
Liker, J.K., *The Toyota Way: 14 Management Principles from the World's Greatest Manufacturer*, McGraw-Hill, New York, 2004.
Meeusen, W. and van den Broeck, J., Efficiency estimation from Cobb–Douglas production functions with composed error, *Int. Econ. Rev.*, 18 (2), 435–444, 1977.
Nahmias, S., *Production and Operations Analysis*, 6th ed., McGraw-Hill, New York, 2009.
Nof, S.Y., Wilhelm, W.E., and Warnecke, H.-J., *Industrial Assembly*, Chapman & Hall, London, 1997.
Ohno, T., *Toyota Production System: Beyond Large-Scale Production*, Productivity Press Inc., Cambridge, MA, 1988a.
Ohno, T., *Just-in-Time for Today and Tomorrow*, Productivity Press Inc., Cambridge, MA, 1988b.
Richmond, J., Estimating the efficiency of production, *Int. Econ. Rev.*, 15, 515–521, 1974.
Semiconductor Equipment and Material International, *Standard for Definition and Measurement of Equipment Productivity*, SEMI E79-0200, Cambridge, MA, 2000.
Semiconductor Equipment and Material International, *Standard for Definition and Measurement of Equipment Reliability, Availability, and Maintainability*, SEMI E10-0701, 2001.
Smith, W., *Time Out: Using Visible Pull Systems to Drive Process Improvements*, John Wiley & Sons, New York, 1998.
Womack, J.P. and Jones, D.T., *Lean Thinking: Banish Waste and Create Wealth in Your Corporation*, 2nd ed., Free Press Simon & Schuster, New York, 2003.

chapter three

Industrial Revolution, customers, and process improvement

Alan R. Heminger

Contents

3.1 Introduction

Over the past few decades, the process approach has come to dominate our view of how to conceptualize and organize work. Current approaches to management, such as business process reengineering (BPR) (Hammer and Champy, 1993), Lean (Womack and Jones, 2003), and Six Sigma (Pande et al., 2000), are all based on this concept. Indeed, it seems almost axiomatic today to assume that this is the correct way to understand organizational work. Yet, each of these approaches seems to say different things about processes. What do they have in common that supports using a process approach? And what do their different approaches tell us about different types of problems with the management of organizational work? To answer these questions, it may help to take a historical look at how work has been done since before the Industrial Revolution up to today.

Before the Industrial Revolution, work was done largely by craftsmen, who underwent a process of becoming skilled in their trade of satisfying customers' wants and needs. Typically, they started as apprentices, where they learned the rudiments of their craft from beginning to end, moved on to become journeymen, then craftsmen as they become more knowledgeable, and finally reaching the pinnacle of their craft as master craftsmen. They grew both in knowledge of their craft and in understanding what their customers wanted. In such an arrangement, organizational complexity was low, with a few journeymen and apprentices working for a master craftsman. However, because work by craftsmen was slow and labor intensive, only a few of the very wealthiest people could have their needs for goods met. Most people did not have access to the goods that the few at the top of the economic ladder were able to get. There was a long-standing and persistent unmet demand for more goods.

This unmet demand, coupled with a growing technological capability, provided the foundations for the Industrial Revolution. Manufacturers developed what Adam Smith (1776) called the "division of labor," in which complex tasks were broken down into simple tasks, automated where possible, and supervisors/managers were put in place to see that the pieces came together as a finished product. As we moved further into the Industrial Revolution, we continued to increase our productivity and the complexity of our factories. With the huge backlog of unmet demand, there was a willing customer for most of what was made. But, as we did this, an important change was taking place in how we made things. Instead of having a master craftsman in charge who knew both how to make goods as well as what the customers wanted and needed, we had factory supervisors who learned how to make the various parts of the manufactured goods come together. Attention and focus began to turn inward from the customers to the process of monitoring and supervising complex factory work.

Over time, our factories became larger and ever more complex. More and more management attention needed to be focused inward on the issues of managing this complexity to turn out ever higher quantities of goods. In the early years of the 20th century, Alfred Sloan, at General Motors, did for management what the Industrial Revolution had done for labor. He broke management down into small pieces, and assigned authority and responsibility tailored to those pieces. This allowed managers to focus on small segments of the larger organization, and to manage according to the authority and responsibility assigned. Through this method, General Motors was able to further advance productivity in the workplace. Drucker (1993) credits this internal focus on improved productivity for the creation of the middle class over the past 100 years. Again, because of the long-standing unmet demand, the operative concept was that if you could make it, you could sell it. The ability to turn out huge quantities of goods culminated in the vast quantities of goods created in the United States during and immediately following World War II. This was added to by manufacturers in other countries, which came back on line after having their factories damaged or destroyed by the effects of the war. As they rebuilt and began producing again, they added to the total quantities of goods being produced.

Then, something happened that changed everything. Supply started to outstrip demand. It did not happen everywhere evenly, either geographically, or by industry. But, in ever-increasing occurrences, factories found themselves supplying more than people were demanding. We had reached a tipping point. We went from a world where demand outpaced supply to a world where more and more, supply outpaced demand (Hammer and Champy, 1993). Not everything being made was going to sell; at least not for a profit. When supply outstrips demand, customers can choose. And when customers can choose, they will choose. Suddenly, manufacturers were faced with what Hammer and Champy call the "3 Cs": customers, competition, and change (Hammer and Champy, 1993). Customers were choosing among competing products, in a world of constant technological change. To remain in business, it was now necessary to produce those products that customers will choose. This required knowing what customers wanted. However, management and the structure of organizations from the beginning of the Industrial Revolution had been largely focused inward, on raising productivity and making more goods for sale. Managerial structure, information flows, and decision points were largely designed to support the efficient manufacturing of more goods, not on tailoring productivity to the needs of choosy customers.

3.2 Business process reengineering

A concept was needed that would help organizations focus on their customers and their customers' needs. A process view of work provided a path for refocusing organizational efforts on meeting customer needs and expectations. On one level, a process is simply a series of steps, taken in some order, to achieve some result. Hammer and Champy, however, provided an important distinction in their definition of a process. They defined it as "a collection of activities that takes one or more inputs and creates an output that is of value to the customer (1993)." By adding the customer to the definition, Hammer and Champy provided a focus back on the customer, where it had been before the Industrial Revolution. In their 1993 book, *Reengineering the Corporation: A Manifesto for Business Revolution*, Hammer and Champy advocated business process reengineering (BPR), which they defined as "the fundamental rethinking and radical redesign of business processes to achieve dramatic improvements in critical, contemporary measures of performance .... " In that definition, they identified four words they believed were critical to their understanding of reengineering. Those four words were "fundamental," "radical," "dramatic," and "processes." In the following editions of their book, which came out in 2001 and 2003, they revisited this definition, and decided that the key word underlying all of their efforts was the word "process." And with process defined as "taking inputs, and turning them into outputs of value to a customer," customers and what customers' value are the focus of their approach to reengineering.

Hammer and Champy viewed BPR as a means to rethink and redesign organizations to better satisfy their customers. BPR would entail challenging the assumptions under which the organization had been operating, and to redesign around their core processes. They viewed the creative use of information technology as an enabler that would allow them to provide the information capabilities necessary to support their processes while minimizing their functional organizational structure.

3.3 Lean

At roughly the same time that this was being written by Hammer and Champy, Toyota was experiencing increasing success and buyer satisfaction through its use of Lean, which is a process view of work focused on removing waste from the value stream. Womack and Jones (2003) identified the first of the Lean principles as value. And, they state, "Value can only be defined by the ultimate customer." Thus, once again, we see a management concept that leads organizations back to focus on their customers. Lean is all about identifying waste in a value stream (similar to Hammer and Champy's process) and removing that waste wherever possible. But the identification of what is waste can only be determined by what contributes or does not contribute to value, and value can only be determined by the ultimate customer. Therefore, once again, we have a management approach that refocuses organizational work on the customers and their values.

Lean focuses on five basic concepts: value, the value stream, flow, pull, and perfection. "Value," which is determined by the ultimate customer, and the "value stream" can be seen as similar to Hammer and Champy's "process," which focuses on adding value to its customers. "Flow" addresses the passage of items through the value stream, and it strives to maximize the flow of quality production. "Pull" is unique to Lean and is related to the "just-in-time" nature of current manufacturing. It strives to reduce in-process inventory that is often found in large manufacturing operations. "Perfection" is the goal that drives Lean. It is something to be sought after, but never to be achieved. Thus, perfection provides the impetus for constant process improvement.

3.4 Six Sigma

In statistical modeling of manufacturing processes, sigma refers to the number of defects per given number of items created. Six Sigma refers to a statistical expectation of

3.4 defects per million items. General Electric adopted this concept in the development of the Six Sigma management strategy in 1986. While statistical process control can be at the heart of a Six Sigma program, General Electric and others have broadened its use to include other types of error reduction as well. In essence, Six Sigma is a program focused on reducing errors and defects in an organization. While Six Sigma does not explicitly refer back to the customer for its source of creating quality, it does address the concept of reducing errors and variations in specifications. Specifications can be seen as coming from customer requirements; thus, again, the customer becomes key to success in a Six Sigma environment.

Six Sigma makes the assertion that quality is achieved through continuous efforts to reduce variation in process outputs. It is based on collecting and analyzing data, rather than depending on hunches or guesses as a basis for making decisions. It uses the steps define, measure, analyze, improve, and control (DMAIC) to improve existing processes. To create new processes, it uses the steps define, measure, analyze, design, and verify (DMADV). Unique to this process improvement method, Six Sigma uses a series of karate-like levels (yellow belts, green belts, black belts, and master black belts) to rate practitioners of the concepts in organizations. Many companies who use Six Sigma have been satisfied by the improvements that they have achieved. To the extent that output variability is an issue for quality, it appears that Six Sigma can be a useful path for improving quality.

3.5 Selecting a method

From the above descriptions, it is clear that while each of these approaches uses a process perspective, they address different problem sets, and they suggest different remedies (Table 3.1). BPR addresses the problem of getting a good process for the task at hand. It recognizes that many business processes over the years have been designed with an internal focus, and it uses a focus on the customer as a basis for redesigning processes that explicitly address what customers need and care about. This approach would make sense where organizational processes have become focused on internal management needs, or some other issues, rather than on the needs of the customer.

The Lean method came out of the automotive world, and is focused on gaining efficiencies in manufacturing. Although it allows for redesigning brand new processes, its focus appears to be most focused on working with an existing assembly line and finding ways to reduce its inefficiencies. This approach would make sense for organizations that have established processes/value streams where there is a goal to make those processes/value streams more efficient.

Six Sigma was developed from a perspective of statistical control of industrial processes. At its heart, it focuses on variability in processes and error rates in production and

Table 3.1 Process Improvement Methods and Their Areas of Focus

Method	Areas addressed	Solution set
Business Process Reengineering (BPR)	Ineffective, inefficient processes	Create a better process, typically by radical redesign
Lean	Waste in the value stream	Identify wasted steps in the value stream, and where possible eliminate them
Six Sigma	Errors and variability of outputs	Identify causes of errors and variable outputs, often using statistical control techniques, and find ways to control for them

seeks to control and limit variability and errors where possible. It asserts that variability and errors cost a company money, and learning to reduce these will increase profits. Similar to both BPR and Lean, it is dependent on top level support to make the changes that will provide its benefits.

Whichever of these methods is selected to provide a more effective and efficient approach to doing business, it may be important to remember the lessons of the history of work since the beginning of the Industrial Revolution. We started with craftsmen satisfying the needs of a small base of customers. We then learned to increase productivity to satisfy the unmet demand of a much larger customer base, but in organizations that were focused inward on issues of productivity, not outward toward the customers. Now that we have reached a tipping point where supply can overtake demand, we need to again pay attention to customer needs for our organizations to survive and prosper (Gerstner 2002). One of the process views of work may provide the means to do that.

References

Drucker, P.F., *The Post Capitalist Society*, Harper Collins, New York, 1993.

Hammer, M. and Champy, J., *Reengineering the Corporation: A Manifesto for Business Revolution*, Harper Business, New York, 1993.

Gerstner, L., *Who Says Elephants Can't Dance?*, Harper, New York, 2002.

Pande, P.S., Neuman, R.P., and Cavanaugh, R.R., *The Six Sigma Way: How GE, Motorola, and Other Top Companies Are Honing Their Performance*, McGraw-Hill, New York, 2000.

Smith, A., *The Wealth of Nations*, Simon & Brown, 1776.

Womack, J.P. and Jones, D.T., *Lean Thinking: Banish Waste and Create Wealth in Your Corporation*, Free Press, New York, 2003.

chapter four

Performance measurement system for value improvement of services

Michihiro Amagasa

Contents

4.1 Introduction

In today's competitive business situations characterized by globalization, short product life cycles, open systems architecture, and diverse customer preferences, many managerial innovations such as the just-in-time inventory management, total quality management, Six Sigma quality, customer–supplier partnership, business process reengineering, and supply chain integration, have been developed. Value improvement of services based on value engineering and systems approach (Miles, 1984) is also considered a method of managerial innovation. It is indispensable for corporations to expedite the value improvement of services and provide fine products satisfying the required function with reasonable costs.

This chapter provides a performance measurement system (PMS) for the value improvement of services, which is considered an ill-defined problem with uncertainty (Terano, 1985). To recognize a phenomenon as a problem and then solve it, it will be necessary to grasp the essence (real substance) of the problem. In particular, for the value improvement problems discussed in this chapter, they can be defined as complicated, ill-defined problems since uncertainty in the views and experiences of decision makers, called "fuzziness," is present.

Building the method involves the following processes: (a) selecting measures and building a system recognition process for management problems, and (b) providing the performance measurement system for the value improvement of services based on the system recognition process. We call (a) and (b) the PMS design process, also considered a core decision-making process, because in the design process, strategy and vision are exactly interpreted, articulated with, and translated into a set of qualitative and/or quantitative measures under the "means to purpose" relationship.

We propose in this chapter a system recognition process that is based on system definition, system analysis, and system synthesis to clarify the essence of the ill-defined problem. Further, we propose and examine a PMS based on the system recognition process as a value improvement method for services, in which the system recognition process reflects the views of decision makers and enables one to compute the value indices for the resources. In the proposed system, we apply the fuzzy structural modeling for building the structural model of PMS. We introduce the fuzzy Choquet integral to obtain the total value index for services by drawing an inference for individual linkages between the scores of PMS, logically and analytically. In consequence, the system we suggest provides decision makers with a mechanism to incorporate subjective understanding or insight about the evaluation process, and also offers a flexible support for changes in the business environment or organizational structure.

A practical example is illustrated to show how the system works, and its effectiveness is examined.

4.2 System recognition process

Management systems are considered to include cover for large-scale complicated problems. However, for a decision maker, it is difficult to know where to start solving ill-defined problems involving uncertainty.

In general, the problem is classified broadly into two categories. One is a problem with preferable conditions—the so-called well-defined problem (structured or programmable), which has an appropriate algorithm to solve it. The other one is a problem with non-preferable conditions—the so-called ill-defined problem (unstructured or nonprogrammable), which may not have an existing algorithm to solve it or there may be only a partial algorithm. Problems involving human decision making or large-scale problems with a complicated nature are applicable to that case. Therefore, uncertainties such as fuzziness (ambiguity in decision making) and randomness (uncertainty of the probability of an event) characterize the ill-defined problem.

In this chapter, the definition of management problems is extended to semistructured and/or unstructured decision-making problems (Simon, 1977; Anthony, 1965; Gorry and Morton, 1971; Sprague and Carlson, 1982). It is extremely important to consider the way to recognize the essence of an "object" when necessary to solve some problems in the fields of social science, cultural science, natural science, etc.

This section will give a systems approach to the problem to find a preliminary way to propose the PMS for value improvement of services. In this approach, the three steps taken in natural recognition pointed out by Taketani (1968) are generally applied to the process of recognition development. These steps—phenomenal, substantial, and essential—regarding system recognition are necessary processes to go through to recognize the object.

With the definitions and the concept of systems thinking, a conceptual diagram of system recognition can be described as in Figure 4.1.

The conceptual diagram of system recognition will play an important role to the practical design and development of the value improvement system for services. Phase 1, phase 2, and phase 3 in Figure 4.1 correspond to the respective three steps of natural recognition described above. At the phenomenal stage (phase 1), we assume that there exists a management system as an object; for example, suppose a management problem concerning

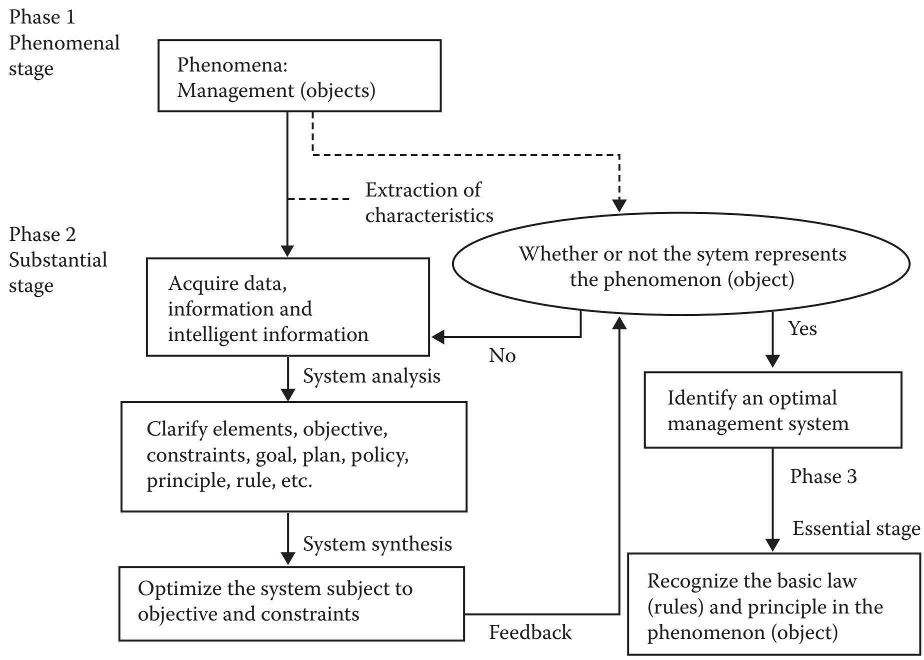

Figure 4.1 Conceptual diagram of system recognition process.

management strategy, human resource, etc., and then extract the characteristics of the problem. Then, in the substantial stage, we may recognize the characteristics of the problem as available information, which are extracted at the previous step, and we perform systems analysis to clarify the elements, objective, constraints, goal, plan, policy, principle, etc., concerning the problem. Next, the objective of the problem is optimized subject to constraints arising from the viewpoint of systems synthesis so that the optimal management system can be obtained. The result of the optimization process, as feedback information, may be returned to phase 1 if necessary, comparing with the phenomena at stage 1.

The decision maker examines whether the result will meet the management system he conceives in his mind (mental model). If the result meets the management system conceived in the phenomenal stage, it becomes the optimal management system and proceeds to the essential stage (phase 3). The essential stage is considered a step to recognize the basic laws (rules) and principles residing in the object. Otherwise, going back to the substantial stage becomes necessary, and the procedure is continued until the optimal system is obtained.

4.3 PMS for value improvement of services

A PMS should act flexibly in compliance with changes in social and/or business environments. In this section, a PMS for the value improvement of services is suggested as shown in Figure 4.2.

At stage A, the algorithm starts at the initial stage, termed structural modeling, in which each model of the function and the cost with respect to services is built up in its own way through the processes encircled with the dotted line in Figure 4.1. For obtaining a concrete model for every individual case, we apply the fuzzy structural modeling method (FSM) (Tazaki and Amagasa, 1979; Amagasa, 2004) to depict an intuitively

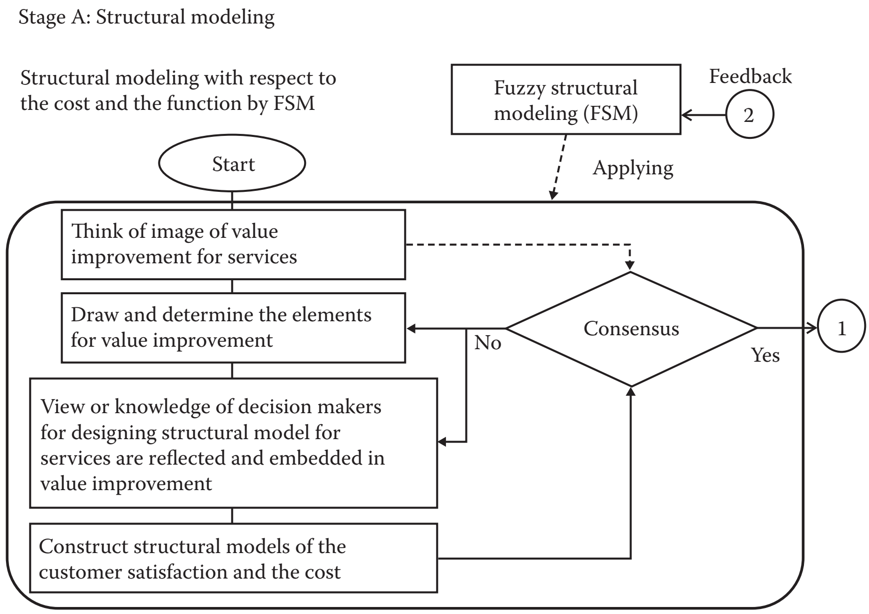

Figure 4.2 Performance measurement system for value improvement of services (Stage A).

graphical hierarchy with well-preserved contextual relations among measured elements. For FSM, binary fuzzy relation within the closed interval of [0, 1] based on fuzzy set (Zadeh, 1965) is used to represent the subordination relations among the elements, and relaxes the transitivity constraint in contrast to ISM (Interpretive Structural Modeling) (Warfield et al., 1975) or DEMATEL (Decision Making Trial and Evaluation Laboratory) (Gabus and Fontela, 1975). The major advantage of those methods may be found in showing intuitive appeal of the graphical picture to decision makers.

First, the decision makers' mental model (imagination) about the given problem, which is the value improvement of services, is embedded in a subordination matrix and then reflected on a structural model. Here, the measured elements are identified by methods such as nominal group techniques (NGT) (Delbecq et al., 1975, 1995), survey with questionnaire, or interview depending on the operational conditions. Thus, we may apply NGT in extracting the measured elements composing the service value and regulating them, clarifying the measurement elements and the attributes. Then, the contextual relations among the elements are examined and represented on the assumption of "means to purpose." The hierarchy of the measurement system is constructed and regarded as an interpretative structural model. Furthermore, to compare the structural model with the mental model, a feedback for learning will be performed by group members (decision makers). If an agreement among the decision makers is obtained, then the process proceeds to the next stage, and the result is considered to be the outcome of stage A. Otherwise, the modeling process restarts from the embedding process or from drawing out and representing the measurement elements process. Then, the process may continue to make progress in the same way as illustrated in Figure 4.2 until a structural model with some consent is obtained.

Thus, we obtain the models of the function and the cost for services as the outcomes of stage A, which are useful for applying to the value improvement of services. Further, we extract and regulate the functions used to perform the value improvement of services by making use of the NGT method described above.

4.3.1 Structural model of functions composing customer satisfaction

We provide, as shown in Figure 4.3, an example of a structural model of function showing the relations between elements (functions) used to find the value of services, which is identified by making use of FSM. In this example, customer satisfaction consists of a set of service functions such as "employee's behavior," "management of a store," "providing customers with information," "response to customers," "exchange of information," and

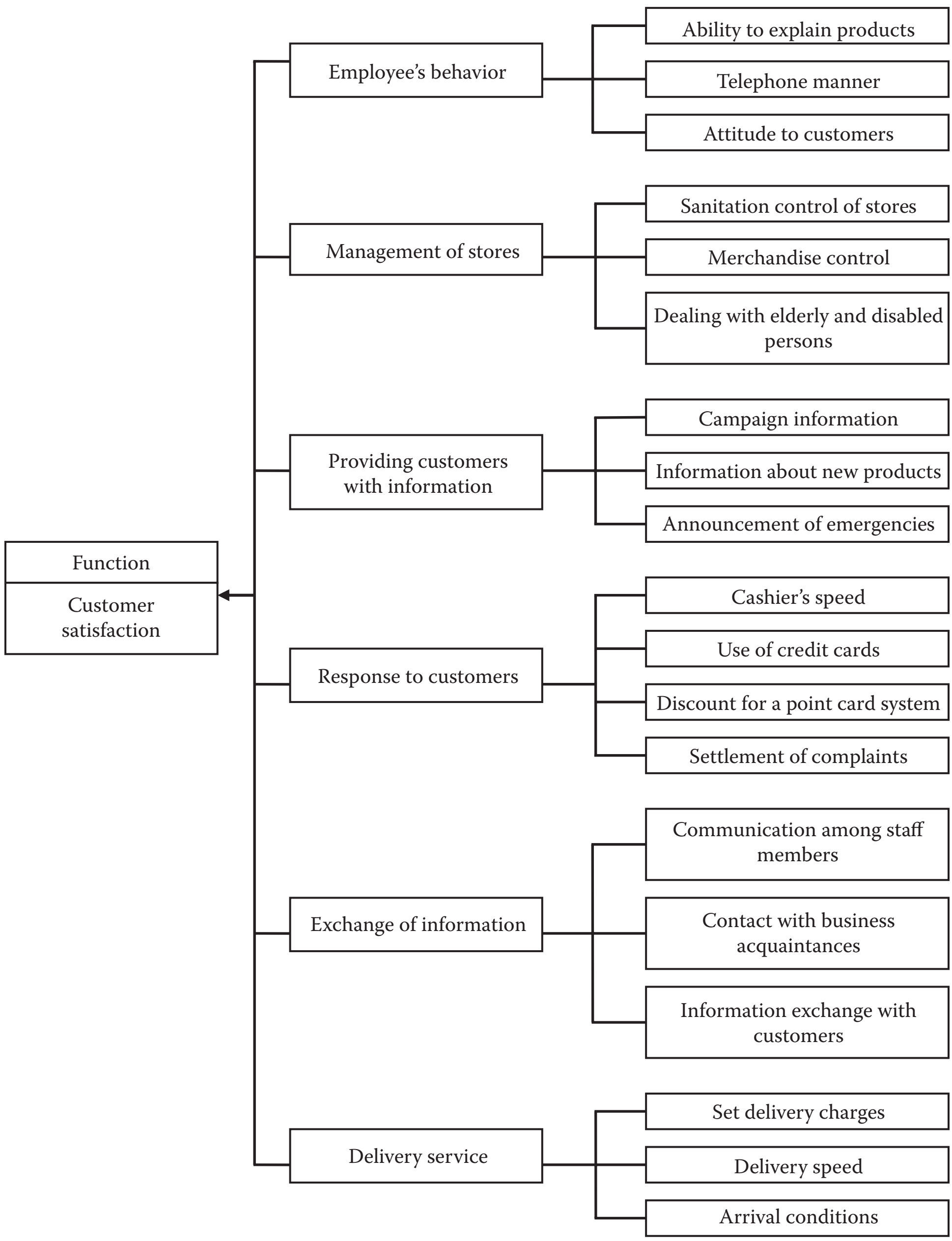

Figure 4.3 Example of structural model of customer satisfaction.

"delivery service." In addition, for each function, "employee's behavior" is described as functions such as "ability to explain products," "telephone manner," and "attitude toward customers." For "management of stores," "sanitation control of stores," "merchandise control," and "dealing with elderly and disabled persons" are enumerated. "Providing customers with information" includes "campaign information," "information about new products," and "announcement of emergencies." "Response to customers" consists of "cashier's speed," "use of credit cards," "discount for a point card system," and "settlement of complaints." In "exchange of information," "communication among staff members," "contact with business acquaintances," and "information exchange with customers" are included. Finally, "delivery service" contains some functions of "set delivery charges," "delivery speed," and "arrival conditions."

4.3.2 Structural model of resources composing cost

Resources (subcosts) composing the cost are also extracted and regulated with the NGT method. An example is illustrated in Figure 4.4 to show the structural model with some resources (subcosts) constituting the cost that is used to offer services in this chapter. Resource (cost) consists of "human resources," "material resources," "financial resources," and "information resources," each of which is also identified by using FSM in the same way as the customer satisfaction was identified. Furthermore, costs relevant to human resources

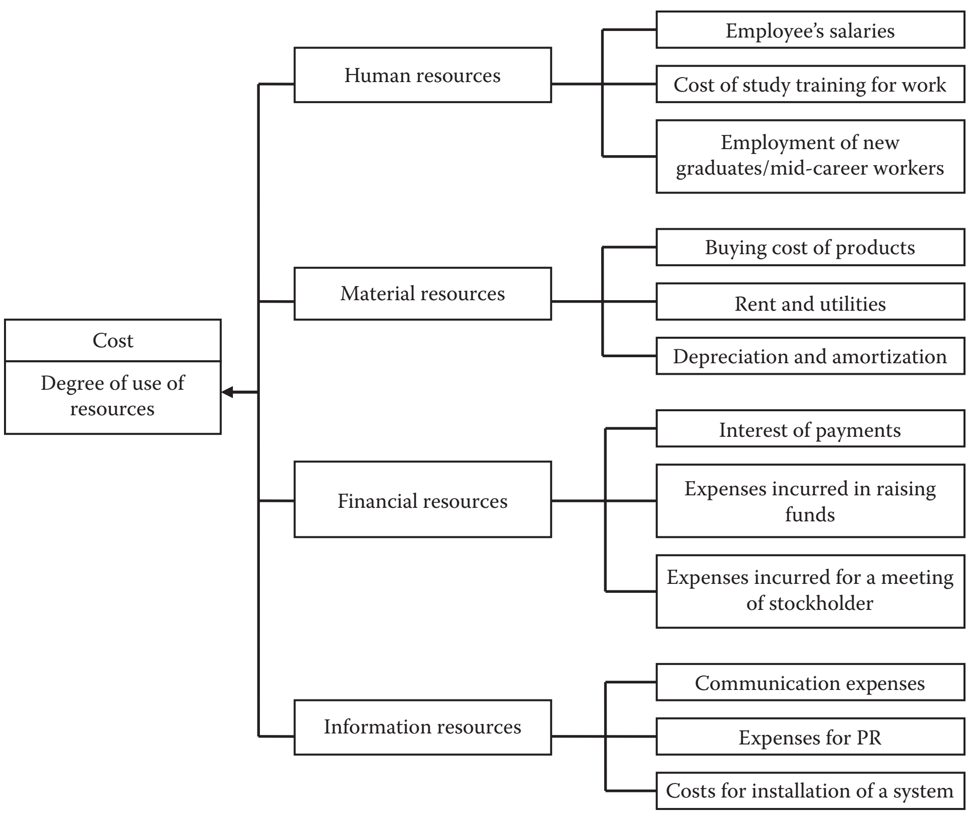

Figure 4.4 Example of structural model of cost.

consist of "employee's salaries," "cost of study training for work," and "employment of new graduates/mid-career workers." "Material resources" contain some subcosts such as "buying cost of products," "rent and utilities," and "depreciation and amortization." "Financial resources" consists of subcosts that are "interest of payments," "expenses incurred in raising funds," and "expenses incurred for a meeting of stockholders." Subcosts for "information resources" are "communication expenses," "expenses for PR," and "costs for installation of a system."

With the structural models of customer satisfaction and the resources (costs) mentioned above, we evaluate the value indices of services.

At stage B shown in Figure 4.2, the value indices for use of four resources, which consist of human resources (R_1), material resources (R_2), financial resources (R_3), and information resources (R_4), are evaluated on the basis of the structural models identified at stage A to perform the value improvement of services.

The weights can be computed by using the Frobenius theorem or the ratio approach with transitive law (Furuya, 1957; Amagasa and Cui, 2009). In this chapter, we use the ratio approach to compute the weights of the function and the cost in the structural models shown in Figures 4.3 and 4.4, and their weights are also used in multi-attribute decision making.

4.3.2.1 Ratio method

The importance degrees of service functions are computed by using the ratio between the functions as follows:

Let F be a matrix determined by paired comparisons among the functions.

Assume that reflexive law is not satisfied in F, and only each element corresponding to $f_{i,i+1}$ ($i = 1, 2, \ldots, n-1$) of the matrix is given as an evaluation value,

	f_1	f_2	f_3	•	f_{n-1}	f_n
f_1	0	f_{12}	–	•	–	
f_2	f_{21}	0	f_{23}	•	–	–
f_3	0	f_{32}	0	•	–	–
•				•		
f_{n-1}	–	–	–	•	0	$f_{n-1,n}$
f_n	–	–	–	•	$f_{n,n-1}$	0

where $0 \leqq f_{i,i+1} \leqq 1$ and $f_{i+1,i}$ satisfies the relation $f_{i+1,i} = 1 - f_{i,i+1}$ ($i = 1, 2, \ldots, k, \ldots, n-1$).

Then, the weight vector $E(=\{E_i,\ i = 1, 2, \ldots, n\})$ of functions (F_i, $i = 1, 2, \ldots, n$) can be found below,

$$
\begin{aligned}
E_1 &= \prod_{i=1}^{n-1} f_{i,i+1} \\
E_k &= \prod_{i=1}^{k-1}(1 - f_{i,i+1}) \prod_{i=1}^{n-1} f_{i,i+1} (1 < k < n, \text{ integer}) \\
E_n &= \prod_{i=1}^{n-1}(1 - f_{i,i+1})
\end{aligned}
\tag{4.1}
$$

We apply the formulas mentioned above to find the weights of functions used. Then, the matrix is constituted with paired comparisons by decision makers (specialists) who take part in the value improvement of services in the corporation. Figure 4.5 shows stages B and C of the PMS.

(1) Importance degree of functions composing customer satisfaction

Suppose, in this chapter, that the functions composing customer satisfaction are extracted and regulated as a set as follows:

$$F = \{F_i, i = 1, 2, \ldots 6\}$$

= {Employee's behavior, Management of a store,
Providing customers with information, Response to customers,
Exchange of information, Delivery service}

Improvement of customer satisfaction becomes a main purpose of corporate management, and $F_i(i = 1, 2, \ldots, 6)$ are respectively defined as the function to achieve customer satisfaction.

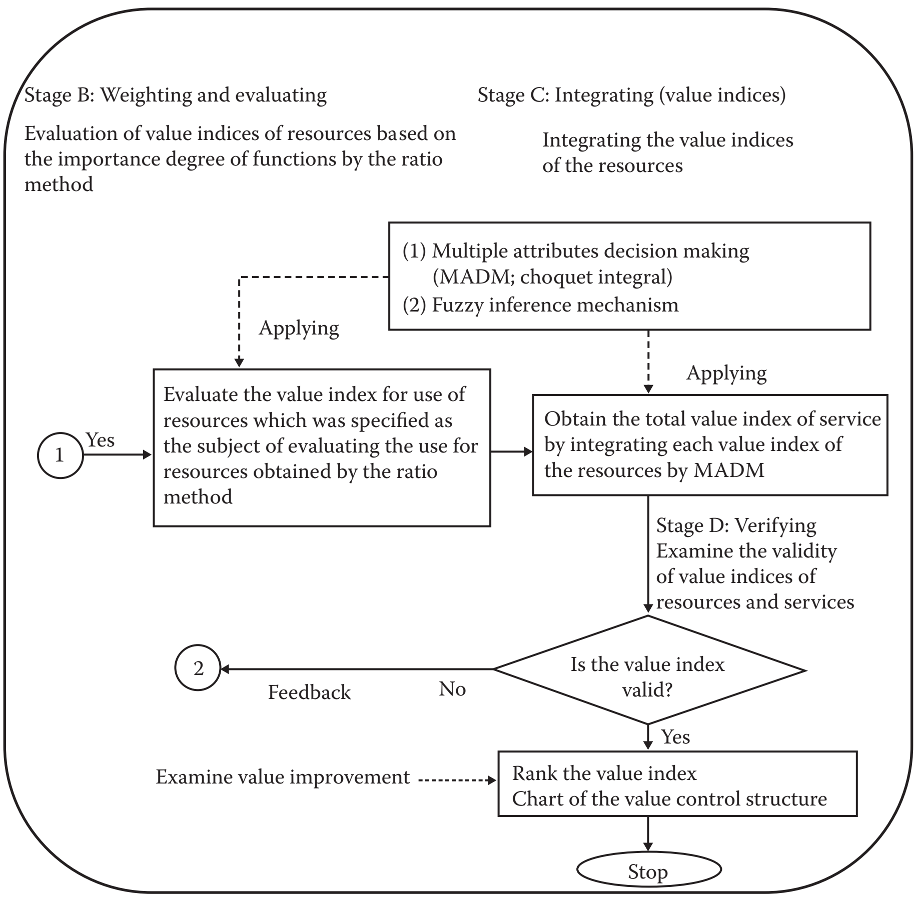

Figure 4.5 Performance measurement system for value improvement of services (Stages B and C).

Then, for example, let each cell of the matrix be intuitively and empirically filled in a paired comparison manner whose values are given by the ratio method by taking into consideration the knowledge and/or experiences of the decision makers (specialists):

	F_1	F_2	F_3	F_4	F_5	F_6
F_1	–	0.8			φ	
F_2		–	0.4			
F_3			–	0.3		
F_4				–	0.6	
F_5					–	0.8
F_6						–

Also, assume that as an evaluation standard to apply paired comparison, we specify five different degrees of grade based on the membership functions.

Not important: [0.0, 0.2)
Not so important: [0.2, 0.4)
Important: [0.4, 0.6)
Very important: [0.6, 0.8)
Critically important: [0.8, 1.0)

For instance, if F_i is believed to be critically more important than F_j, the decision makers may make an entry of 0.9 in F_{ij}. Each value is empirically given by the decision makers (or specialists) who have their experiences and knowledge, with the know-how for value improvement. As a result, the values yielded by the ratio method are recognized as weights for the functions.

Thus, the weight vector E of functions (F_i, $i = 1, 2, \ldots, 6$) is obtained as follows:

$$E = \{0.046, 0.012, 0.017, 0.040, 0.027, 0.007\}$$

Further, F can be standardized

$$E = \{0.31, 0.08, 0.11, 0.27, 0.18, 0.05\}$$

a. *Importance degrees of constituent elements of "employee's behavior (F_1)"*
   i. As it is clear from the structural model of the customer satisfaction shown earlier in Figure 4.3, F_1 consists of all subfunctions F_{1i} ($i = 1, 2, 3$).
   ii. Here, we compute the importance degrees of $\{F_{1i}, i = 1, 2, 3\}$ by the ratio method in the same way as F_1 was obtained.
b. *Importance degrees of subfunctions of "employee's behavior (F_1)"*

$$\begin{aligned} F_1 &= \{F_{1i}, i = 1, 2, 3\} \\ &= \{\text{Explainable ability for products,} \\ &\quad \text{Telephone manner, Attitude toward customers}\} \end{aligned}$$

F_1	F_{11}	F_{12}	F_{13}
F_{11}	–	0.6	
F_{12}		–	0.3
F_{13}			–

i. Then the weight vector $E(=\{E_{1i}, i = 1, 2, 3\})$ for $\{F_{1i}, i = 1, 2, 3\}$ is found as follows:

$$E = \{0.31, 0.21, 0.48\}$$

ii. From this, the importance degrees $\{E_{1i}, i = 1, 2, 3\}$ of subfunctions $\{F_{1i}, i = 1, 2, 3\}$ are also recomputed with weight of F_1 as follows:

$$E_{11} = \text{weight of } F_1 \times \text{weight of } F_{11} = 0.31 \times 0.31 = 0.096$$
$$E_{12} = \text{weight of } F_1 \times \text{weight of } F_{12} = 0.31 \times 0.21 = 0.065$$
$$E_{13} = \text{weight of } F_1 \times \text{weight of } F_{13} = 0.31 \times 0.48 = 0.149$$

iii. In a similar way, the weights of other functions $F_i (i = 2, 3, \ldots, 6)$ and the importance degrees for subfunctions of $F_i (i = 2, 3, \ldots, 6)$ are obtained by the ratio method. The computational results are summarized in Table 4.1.

(2) Amount of the cost (resources) based on the structural model of cost

The cost is understood as the amount of resources utilized to provide the customers with the services. To calculate the cost for services, we prepare the questionnaire

Table 4.1 Weights of Subfunctions to Improve Customer Satisfaction

	Subfunctions	Weights
F_1	F_{11} (Ability to explain products)	0.096
	F_{12} (Telephone manner)	0.065
	F_{13} (Attitude toward customers)	0.149
F_2	F_{21} (Sanitation control of stores)	0.026
	F_{22} (Merchandise control)	0.038
	F_{23} (Dealings with elderly and disabled persons)	0.016
F_3	F_{31} (Campaign information)	0.039
	F_{32} (Information about new products)	0.057
	F_{33} (Announcement of emergencies)	0.014
F_4	F_{41} (Cashier's speed)	0.059
	F_{42} (Use of credit cards)	0.024
	F_{43} (Discount for a point card system)	0.038
	F_{44} (Settlement of complaints)	0.149
F_5	F_{51} (Communication among staff members)	0.056
	F_{52} (Contact with business acquaintances)	0.038
	F_{53} (Information exchange with customers)	0.086
F_6	F_{61} (Set delivery charges)	0.031
	F_{62} (Delivery speed)	0.008
	F_{63} (Arrival conditions)	0.012

for the decision makers (specialists); that is, how many resources were utilized in every possible way to pursue/achieve the value of services?

a. *Evaluation of cost (resources)*
   i. Let us denote C by the amount utilized of four resources. These are expressed by $C_i (i = 1, 2, \ldots, 4)$ as below.

$$\begin{aligned} C &= \{C_i, i = 1, 2, \ldots, 4\} \\ &= \{\text{Human resources, Material resources,} \\ &\quad \text{Financial resources, Information resources}\} \end{aligned}$$

   ii. The degrees for use of resources is meant by the purpose of corporation for using resources effectively, and $C_i (i = 1, 2, \ldots, 4)$ are considered the costs to achieve the purpose.
   iii. The following matrix shows responses provided by the decision makers (specialists) answering the questionnaire.

C	C_1	C_2	C_3	C_4
C_1	–	0.6		
C_2		–	0.7	
C_3			–	0.4
C_4				–

   iv. Applying Equation 4.1 to the matrix, we can obtain the subcosts utilized to give services, that is,

$$\{C_i, i = 1, 2, \ldots, 4\} = \{0.42, 0.28, 0.12, 0.18\}.$$

   v. For instance, "$C_1 = 0.42$" shows the amount of human resources utilized to perform the services.

b. *Evaluation of subcost composing the human resources*

$$\begin{aligned} C_1 &= \{C_{1i}, i = 1, 2, 3\} \\ &= \{\text{Employee's salaries, Cost of study training for work,} \\ &\quad \text{Employment of new graduates/mid-career workers}\} \end{aligned}$$

   i. The following matrix is provided similarly by the decision makers.
   ii. Analogous to the above, we can get the subcosts utilized to give the services.

C_1	C_{11}	C_{12}	C_{13}
C_{11}	–	0.8	
C_{12}		–	0.4
C_{13}			–

$$\{C_{1i}, i = 1, 2, 3\} = \{0.62, 0.16, 0.22\}.$$

iii. Namely, the amount of cost C_1 consists of those of subcosts for human resources.
iv. Then C_{11} = amount of $C_1 \times C_{11} = 0.42 \times 0.62 = 0.26$. "$C_{11} = 0.26$" means the amount of subcost of human resources utilized to give the services.

In a similar way, the amounts of other resources $\{C_i, i = 2, 3, 4\}$ as well as subcosts of $\{C_i, i = 2, 3, 4\}$ are also computed by the ratio method. The computational result is found in Table 4.2, which shows the subcosts for the resources utilized to give the services.

In Table 4.3, a_{ij} shows the degrees of resource R_i used to satisfy the function item F_j, $j = 1, 2, \ldots, m$. $RF_{ij} = E_i \times a_{ij} \times 10^{-2}$, $(j = 1, 2, \ldots, 6)$, $\sum_{j=1}^{6} a_{ij} = 100(\%)$, $(i = 1, 2, 3, 4)$. $E_j (j = 1, 2, \ldots, 6)$ shows the degree of importance of each function items. The costs of the resource $R_i (i = 1, 2, 3, 4)$ will be computed and shown as $\sum_{k=1}^{6} RE_{ik}$ in Table 4.3.

4.3.3 Computing for value indices of four resources

In general, the value index of object in value engineering is defined by the following formula.

$$\text{Value index} = \text{satisfaction for necessity/use of resources} \tag{4.2}$$

The value index is interpreted to show the degree of satisfaction to fill necessity, which is brought by the resources when they are utilized. On the basis of this formula, in this study, we define the value of services composing four resources as below.

$$\text{Value of services} = \text{function of services/cost of services} \tag{4.3}$$

Table 4.2 Weights of Subresources Composing Cost of Services

	Subresources composing cost	Subcosts
C_1	C_{11} (Employee's salaries)	0.260
	C_{12} (Cost of study training for work)	0.067
	C_{13} (Employment of new graduates/mid-career workers)	0.092
C_2	C_{21} (Buying cost of products)	0.162
	C_{22} (Rent and utilities)	0.070
	C_{23} (Depreciation and amortization)	0.048
C_3	C_{31} (Interest of payments)	0.028
	C_{32} (Expenses incurred in raising funds)	0.065
	C_{33} (Expenses for meetings for stockholders)	0.028
C_4	C_{41} (Communication expenses)	0.027
	C_{42} (Expenses for PR)	0.108
	C_{43} (Costs for installation of a system)	0.045

Table 4.3 Importance Degrees of Resources from Functions of Customer Satisfaction

	Function items						
Resources	F_1	F_2	F_3	F_4	F_5	F_6	$\sum_{k=1}^{6} E_k = 100\%$
R_1	$a_{11}\ RE_{11}$	$a_{12}\ RE_{12}$	$a_{13}\ RE_{13}$	$a_{14}\ RE_{14}$	$a_{15}\ RE_{15}$	$a_{16}\ RE_{16}$	$\sum_{k=1}^{6} E_{1k}$
R_2	$a_{21}\ RE_{21}$	$a_{22}\ RE_{22}$	$a_{23}\ RE_{23}$	$a_{24}\ RE_{24}$	$a_{25}\ RE_{25}$	$a_{26}\ RE_{26}$	$\sum_{k=1}^{6} E_{2k}$
R_3	$a_{31}\ RE_{31}$	$a_{32}\ RE_{32}$	$a_{33}\ RE_{33}$	$a_{34}\ RE_{34}$	$a_{35}\ RE_{35}$	$a_{36}\ RE_{36}$	$\sum_{k=1}^{6} E_{3k}$
R_4	$a_{41}\ RE_{41}$	$a_{42}\ RE_{42}$	$a_{43}\ RE_{43}$	$a_{44}\ RE_{44}$	$a_{45}\ RE_{45}$	$a_{46}\ RE_{46}$	$\sum_{k=1}^{6} E_{4k}$

Therefore, the value index, which is based on importance degree and cost concerning each resources used to give services, is obtained.

$$
\begin{aligned}
\text{Value index of human resources} &= \sum_{k=1}^{m} E_{1k} / \text{cost of human resources} \\
\text{Value index of material resources} &= \sum_{k=1}^{m} E_{2k} / \text{cost of material resources} \\
\text{Value index of financial resources} &= \sum_{k=1}^{m} E_{3k} / \text{cost of financial resources} \\
\text{Value index of information resources} &= \sum_{k=1}^{m} E_{4k} / \text{cost of information resources}
\end{aligned}
\tag{4.4}
$$

At stage C, the multi-attribute decision-making method (MADM) based on Choquet integral (Grabisch, 1995; Modave and Grabisch, 1998) can be introduced and a total value index of services (service value) is found by integrating the value indices of the human, material, financial, and information resources. Let X_i ($i = 1, 2$) be fuzzy sets of universe of discourse X. Then the λ fuzzy measure g of the union of these fuzzy set, $X_1 \bigcup X_2$ can be defined as follows:

$$g(X_1 \bigcup X_2) = g(X_1) + g(X_2) + \lambda g(X_1)\, g(X_2)$$

where λ is a parameter with values $-1 < \lambda < \infty$, and note that $g(\cdot)$ becomes identical to probability measure when $\lambda = 0$. Here, since it is assumed that when the assessment of corporation is considered, the correlations between factors are usually independent, the fuzzy sets X_1 and X_2 are independent, that is, $\lambda = 0$. Then, the total value index of services is expressed as in Equation 4.5.

$$\begin{aligned}
&\text{Total value index of services}\\
&\quad = g(\text{value index of human resources, value index of material resources,}\\
&\qquad \text{value index of financial resources, value index of information resources})\\
&\quad = w_1 \times \text{value index for human resources}\\
&\qquad + w_2 \times \text{value index for material resources}\\
&\qquad + w_3 \times \text{value index for financial resources}\\
&\qquad + w_4 \times \text{value index for information resources}
\end{aligned} \tag{4.5}$$

where $w_i (0 \le w_i \le 1; i = 1, 2, 3, 4)$ are weights for respective resources.

At stage D, if the integrated evaluation value is examined and its validity is shown, the process goes to the final stage (stage E).

At stage E, the integrated value indices of services computed in the previous step are ranked using the fuzzy outranking method (Roy, 1991; Siskos and Oudiz, 1986) and draw the graphic structure of value control (Amagasa, 1986). Then the process terminates.

In this study, each of the value indices of services is represented in the graphic structure of the value control depicted.

4.4 Simulation for value improvement system of services

In this section, we carry out a simulation of the procedure to perform the value improvement system of services and examine the effectiveness of the proposed value improvement system.

Here, as specific services trade, we take up a fictitious household appliance store, DD Company. This store is said to be a representative example providing "a thing and services" to customers. The store sells "things" such as household electrical appliances, which are essential necessities of life and commercial items used in everyday life. In addition, it supplies customer services when customers purchase the "thing" itself. DD Company was established in 1947 and the capital is 19,294 million yen, the yearly turnover is 275,900 million yen, total assets are worth 144,795 million yen, the number of the stores is 703 (the number of franchise stores is 582 on March 31, 2007), and the number of employees is 3401. The store is well known to the customers on the grounds that it would make a difference with other companies, by which the management technique is designed for a customer-oriented style, pursuing customer satisfaction. For example, salespersons have sufficient knowledge about the products they sell and give suitable advice and suggestions according to customers' requirements, which often happens on the sales floor. We conducted a survey for DD Company. The simulation was based on the results of a questionnaire survey and performed by applying the PMS for the value improvement of services shown in Figure 4.2.

4.4.1 Stage A: Structural modeling

Figures 4.3 and 4.4 show the structural models with respect to the functions composing customer satisfaction, and the cost showing the use of resources.

4.4.2 Stage B: Weighting and evaluating

Table 4.2 shows the importance degrees of resources for functions of customer satisfaction, which is obtained by consensus among decision makers (specialists) with know-how deeply related to the value improvement of services.

By Table 4.4, it is understood that the distributed amount for four resources and the real ratios, which are used to attain customer satisfaction related to six functions, are provided with the four resources.

From this, each of the value indices with respect to the respective resources used to supply customer services, for which human resources, material resources, financial resources, and information resources are considered, is obtained by using Tables 4.1 through 4.4 and Equation 4.4.

(1) Value index of Human resources = 45.64/42 (= 1.1)
(2) Value index of Material = 4.08/28 (= 0.14)
(3) Value index of Financial = 13.19/12 (= 1.08)
(4) Value index of Information = 36.37/18 (= 2)

From the value indices for the resources mentioned above, the chart of value control graphic structure is depicted as shown in Figure 4.5. Thus, it may be concluded by Figure 4.6 that the following results with respect to the value improvement of services from the viewpoints of function and cost are ascertained.

(1) In this corporation, there is no need for doing the value improvement related to each of human resources, financial resources, and information resources because three of all four resources are below the curved line, implying a good balance between the cost and the function of services.
(2) For material resources, it will be necessary to exert all possible efforts for the value improvement of the resource because the value index is 0.04, which is much smaller than 1.00.
(3) On the whole, the total value index of services is counted 1.23 as shown below, so that the value indices for four resources are included within the optimal zone of the chart of value control graphic structure shown in Figure 4.5. Therefore, it could be concluded that the corporation may not have to improve the value of services of their organization.

4.4.3 Stage C: Integrating (value indices)

At the integrating stage, MADM based on Choquet integral (Grabish, 1995; Modave and Grabish, 1998) can be introduced for the value improvement of services, and the total value index of services is obtained by integrating the value indices of the four resources as follows:

$$\begin{aligned}\text{Total value index of services} &= w_1 \times 1.1 + w_2 \times 0.14 + w_3 \times 1.08 + w_4 \times 2\\ &= 0.46 \times 1.1 + 0.11 \times 0.14 + 0.17 \times 1.08 + 0.26 \times 2 = 1.23\end{aligned}$$

As a result of the simulation, the value of services of DD Company indicates a considerably high level because the total value index becomes 1.23 (>1.00), which belongs to the optimal region.

Table 4.4 Importance Degrees of Resources for Functions of Customer Satisfaction

		Function (Customer satisfaction)																		
		Employee's behavior (0.31)			Management of stores (0.08)			Providing customers with information (0.11)			Response to customers (0.27)				Exchange of information (0.18)			Delivery services (0.05)			
		Ability to explain products	Telephone manner	Attitude to customers	Sanitation control of stores	Merchandise control	Dealing with elderly and disabled persons	Campaign information	Information about new products	Announcement in emergencies	Cashier's speed	Use of credit cards	Discount for a point card system	Settlement of complaints	Communication among staff members	Contact with business acquaintances	Information exchange with customers	Set delivery charges	Delivery speed	Arrival conditions	Integrated evaluation values
Use of resources		0.096	0.065	0.149	0.026	0.038	0.016	0.039	0.057	0.014	0.059	0.024	0.038	0.015	0.056	0.038	0.086	0.031	0.008	0.012	
Human resources	Employee's salaries	70	70	80	20	20	40			30	30	30	30	80	30	20	20		20	20	45.64 (%)
	Cost of study training for work																				
	Employment of new graduates/mid-career workers	6.72	4.55	11.92	0.52	0.76	0.64			0.42	1.77	0.72	1.14	11.92	1.68	0.76	1.72		0.16	0.24	

Material resources	Buying cost of products		10		10	30	20	20										30			4.08 (%)
	Rent and utilities																				
	Depreciation and amortization		0.65		0.26	1.14	0.32	0.78										0.93			
Financial resources	Interest of payments				20	20	20	20	20	20	30	30	30		30	30	20	20	30	30	13.19 (%)
	Expenses incurred in raising funds																				
	Expenses incurred for a meeting of stockholders				0.52	0.76	0.32	0.78	1.14	0.28	1.77	0.72	1.14		1.68	1.14	1.72	0.62	0.24	0.36	
Information resources	Communication expenses	30	20	20	50	30	40	60	60	50	40	40	40	20	40	50	60	50	50	50	36.37 (%)
	Expenses for PR																				
	Costs for installation of a system	2.88	1.3	2.98	1.3	1.14	0.64	2.34	3.42	0.7	2.36	0.96	1.52	2.98	2.24	1.9	5.16	1.55	0.4	0.6	

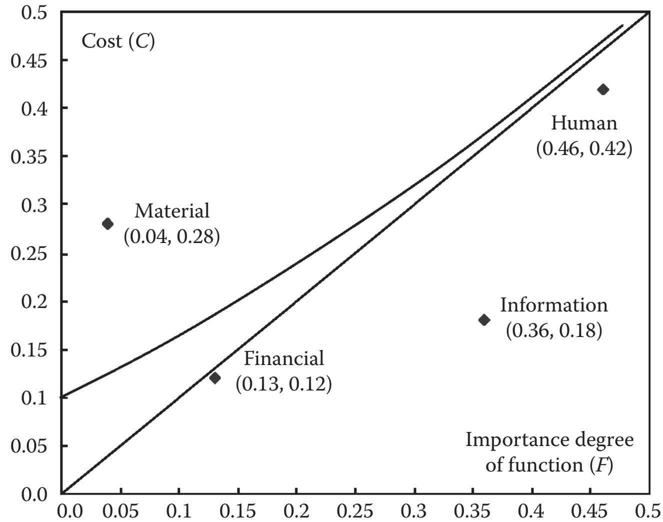

Figure 4.6 Value control graphic structure.

About DD Company, Nikkei Business announces that the store scored high on the evaluation. The store advocates that customers "have trust and satisfaction with buying the products all the time." Also, the store supplies "attractive goods" at "a reasonable price" as well as "superior services," as a household appliance retail trade based on this management philosophy. Furthermore, the store realizes a customer-oriented and community-oriented business, and supplies smooth services reflecting area features and scales advantages by controlling the total stock in the whole group. From this, it can be said that the validity of the proposed method was verified by the result of this simulation experiment, which corresponds to the high assessment of DD Company by Nikkei Business, as described above.

4.5 Conclusion

It is important for an administrative action to pursue profit of a corporation by making effective use of four resources—capable persons, materials, capital, and information. In addition, allowing each employee to attach great importance to services, and then hoping that the employee would willingly improve service quality, and thus enhancing the degree of customer satisfaction, is important in the services trade. These surely promise to bring about profit improvement for the corporation.

We proposed in this chapter a system recognition process that is based on system definition, system analysis, and system synthesis, clarifying the "essence" of an ill-defined problem. Further, we suggest the PMS as a method for the value improvement of services and examined it, in which the system recognition process reflects the views of decision makers and enables to compute the effective service scores. As an illustrative example, we took up the evaluation problem of a household appliance store selected from the viewpoint of service functions, and come up with a new value improvement method by which the value indices of services are computed. To verify the effectiveness of the new method we suggested, we performed a questionnaire survey about the service functions for the household appliance store. As a result, it was determined that the proposed method is significant for the value improvement of services in corporations.

Finally, the soundness of this system was verified by the result of this simulation. With this procedure, it is possible to build PMS for services that is based on realities. This part of the study remains a future subject.

References

Amagasa, M., *System Synthesis Theory* (in Japanese), Moriyama Publishing Co., Tokyo, 1986.

Amagasa, M., *Management Systems Engineering*, Institute of Business Research, Daito-Bunka University, Tokyo, 2004.

Amagasa, M. and Cui, D., *Thinking from the Aspect of Management System* (in Japanese), Souseisya Publishing Co., Tokyo, 2009.

Anthony, R.N., *Planning and Control Systems: A Frame work for Analysis*, Graduate School of Business Administration, Harvard University, Cambridge, MA, 1965.

Delbecq, A.L., Vande Ven, A.H., and Gustafson, H.D., *Group Techniques for Program Planning—A Guide to Nominal Group and Delphi Processes*, Scott, Foresman and Company, San Francisco, 1995.

Delbecq, A.L. et al., *Group Techniques for Program Planning*, Scott, Foresman and Company, San Francisco, 1975.

Furuya, S., *Matrix and Determinant* (in Japanese), New Mathematics Series 5, Baifukan, Tokyo, 1957.

Gabus, A. and Fontela, E., *DEMATEL Reports*, Battelle Geneva Research Centre, Geneva, 1975.

Gorry, G.A. and Morton, M.S.S., A framework for management information system, *Sloan Manage. Rev.*, Fall, 55–70, 1971.

Hogg, M.A. and Tindale, R. Scott. *Blackwell Handbook of Social Psychology: Group Processes*, Blackwell Publisher, pp.107–138, 2001.

Grabisch, M., Fuzzy integral in multicriteria decision making, *Fuzzy Sets Syst.*, 69, 279–298, 1995.

Miles, L.D., *Techniques for Value Analysis and Engineering*, 3rd ed., McGraw-Hill, New York, 1984.

Modave, F. and Grabisch, M., Preference representation by the Choquet integral: The commensurability hypothesis, *Proceedings of the 7th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, Paris, France, 164–171, 1998.

Roy, B., The outranking approach and the foundations of electro methods, *Theory Decis.*, 31, 49–73, 1991.

Simon, H.A., *The New Science of Management Decision*, Harper & Row, New York, 1977.

Siskos, J.J., Lombard, J., and Oudiz, A., The use of multicriteria outranking methods in the comparison of control options against a chemical pollutant, *Oper. Res. Soc.*, 37, 357–371, 1986.

Sprague, Jr., R. and Carlson, E., *Building Effective Decision Support Systems*, Prentice Hall, New Jersey, 1982.

Tazaki, E. and Amagasa, M., Structural modelling in a class of systems using fuzzy sets theory, *Int. J. Fuzzy Sets Syst.*, 2 (1), 87–103, 1979.

Taketani, M., *Various Problems of Dialectic* (in Japanese), Keisosyobo, Tokyo, 1968.

Terano, T., *Introduction to Systems Engineering* (in Japanese), Kyoritsu Publish, Tokyo, 1985.

Warfield, J.N. et al., *A Unified Systems Engineering Concept*, Battelle M. Institute, Ohio, 1975.

Zadeh, L.A., Fuzzy set, *Inf. Control*, 8, 338–353, 1965.

chapter five

Strategic performance measurement

Garry D. Coleman and Altyn Clark

Contents

5.1 What is strategic performance measurement?

The focus of this chapter is strategic performance measurement, a key management system for performing the study (or check) function of Shewhart's Plan–Do–Study–Act cycle. Strategic performance measurement applies to a higher-level system of interest (unit of analysis) and a longer-term horizon than operational performance measurement. While the dividing line between these two types of performance measurement is not crystal clear, the following distinctions can be made:

- Strategic performance measurement applies to the organizational level, whether of a corporation, a business unit, a plant, or a department. Operational performance measurement applies to small groups or individuals, such as a work group, an assembly line, or a single employee.
- Strategic performance measurement is primarily concerned with performance that has medium- to long-term consequences; thus, performance is measured and reported on a weekly, monthly, quarterly, or annual basis. More frequent, even daily, measurement and reporting may also be included, but only for the most important performance measures. Data may also be collected daily or perhaps continually, but should be aggregated and reported weekly or monthly. Operational performance measurement focuses on immediate performance, with reporting on a continual, hourly, shift, or daily basis. Strategic performance measurement tends to measure performance on a periodic basis, while operational performance measurement tends to measure on a continual or even continuous basis.
- Strategic performance measurement is concerned with measuring the mission- or strategy-critical activities and results of an organization. These activities and results are keys to the organization's success, and their measurements are referred to as strategic performance measures, key performance indicators, or mission-driven metrics. These measurements can be classified into a few key performance dimensions, such as Drucker's (1954) nine key results areas, the Balanced Scorecard's four performance perspectives (Kaplan and Norton, 1996), the Baldrige criteria's five business results items (Baldrige Performance Excellence Program, 2013), or Sink's (1985) seven performance criteria.
- Strategic performance measurement tends to measure aspects of performance affecting the entire organization, while operational performance may be focused on a single product or service (out of many). In an organization with only one product, strategic and operational measurement may be similar. In an organization with multiple products or services, strategic performance measurement is likely to aggregate performance data from multiple operational sources.
- Strategic performance measurement is a popular topic in the management, accounting, industrial engineering, human resources management, information technology, statistics, and industrial and organizational psychology literature. Authors such as Bititci et al. (2012), Brown (1996, 2000), Busi and Bititci (2006), Kaplan and Norton (1992, 1996), Neely (1999), Thor (1998), and Wheeler (1993) have documented the need for and the challenges facing strategic performance measurement beyond traditional financial and accounting measures. Operational performance measurement has long been associated with pioneers such as Frederick Taylor, Frank and Lillian Gilbreth, Marvin Mundel, and others. Careful reading of their work often shows an appreciation for and some application to strategic performance measurement, yet they are remembered for their contributions to operational measurement.
- For the remainder of this chapter, strategic performance measurement will be referred to as performance measurement. The term "measurement" will be used to apply to both strategic and operational performance measurement.

Why is performance measurement important enough to warrant a chapter of its own? Andrew Neely (1999, p. 210) summarized the reasons for the current interest in performance measurement very well. His first reason is perhaps the most important for the industrial engineer: the "changing nature of work." As industrialized nations have seen their workforces shift to predominantly knowledge and service work, concerns have arisen about how to measure performance in these enterprises with less tangible products. Fierce competition and a history of measuring performance have facilitated steady productivity and quality improvement in the manufacturing sector in recent years. Productivity and quality improvement in the service sector has generally lagged that of the manufacturing sector. The shift to a knowledge- and service-dominated economy has led to increased interest in finding better ways to measure and then improve performance in these sectors. Other reasons for increased interest in performance measurement cited by Neely include increasing competition, specific improvement initiatives that require a strong measurement component (such as Six Sigma or business process reengineering), national and international awards (with their emphasis on results, information, and analysis), changing organizational roles (e.g., the introduction of the chief information officer or, more recently, the chief knowledge officer), changing external demands (by regulators and shareholders), and the power of information technology (enabling us to measure what was too expensive to measure or analyze in the past).

5.2 Measurement in context of planning

An effective measurement approach enables and aligns individual, group, and organizational Plan–Do–Study–Act spirals to assist people in learning and growth toward a common aim. The Plan–Do–Study–Act spiral permeates human endeavor. Everything people do involves (consciously or unconsciously) four simple steps: make a Plan, Do the plan, Study the results, and Act the same or differently in the future, on the basis of what was learned. Plan–Do is the priority setting and implementation process. Study–Act is the measurement and interpretation process. Study–Act is different than, yet inseparable from, Plan–Do. Plan–Do–Study–Act is a structured and extremely useful (though mechanistic) theory of organizational learning and growth. The essence of Plan–Do–Study–Act within an organization is feedback and learning for the people in the system. Measurement's highest purpose in the context of strategy is to raise group consciousness about some phenomenon in the organization or its environment, thereby enhancing the opportunity to make mindful choices to further organizational aims. A strategic Plan–Do–Study–Act cycle for an organization may be notionally described by asking four fundamental questions: (1) What experiences and results does the organization aim to create over some time horizon? (2) How will people know if or when those experiences and results are occurring? (3) What actions and behaviors are people in the organization committed to, to create those experiences and results? (4) How will people know if those actions and behaviors are occurring? Questions (1) and (3) are strategic planning questions, while (2) and (4) are strategic measurement questions. Answers to questions (1) and (2) generally take the form of desired outcomes: nouns and adjectives. Answers to questions (3) and (4) generally take the form of planned activities: verbs and adverbs. Senior leaders have an obligation to answer questions (1) and (2) to provide direction and communicate expectations for the organization. Senior leaders are a participatory resource to help others in the organization shape answers to questions (3) and (4). One very important (though limited) view of leadership is the leader as organizational hypothesis tester: "If people act and behave question (3) answers—as verified by question (4) indicators—then question (1) results—as measured by question (2) indicators—are more likely to occur." It is this implicit hypothesis testing that links planning and measurement in the management process.

5.3 Measurement and evaluation process

Measurement is a human procedure of using language, images, and numbers to codify feedback from the universe about individual, organizational, and societal effectiveness—the extent, size, capacity, characteristics, quality, and amount or quantity of objects or events. In an organizational setting, measurement is the codifying of observations into data that can be analyzed, portrayed as information, and evaluated to support the decision maker. The term "observation" is used broadly here and may include direct observation by a human, sensing by a machine, or document review. Document review may involve secondary measurement, relying on the recorded observations of another human or machine; or it may involve the direct measurement of some output or artifact contained in the documents. The act of measurement produces data ("evidence"), often but not always in quantified form. Quantitative data are often based on counts of observations (e.g., units, defects, person-hours) or scaling of attributes (e.g., volume, weight, speed). Qualitative data are often based on categorization of observations (e.g., poor/fair/good) or the confirmation (or not) of the presence of desired characteristics (e.g., yes/no, pass/fail). Such qualitative data are easily quantified by calculating the percentages in each category. See Muckler and Seven (1992) for a thoughtful discussion of the related question of objective versus subjective measurement.

Measuring performance—both strategic performance and operational performance—is a process (see Figure 5.1) that produces a codified representation of the phenomenon being measured. Assuming it was measured properly, this codified representation is simply a fact. This fact may exist in the form of a number, chart, picture, or text, and is descriptive of the phenomena being observed (i.e., organizational performance) and the process used to produce the fact before evaluation. Evaluation is the interpretation and judgment of the output of the measurement process (i.e., the number, chart, picture, or text). Evaluation results in a determination of the desirability of the level or trend of performance observed, typically on the basis of a comparison or expectation. Too often, those who are developing new or enhanced performance indicators jump to evaluation before fully completing the measurement step. They base the suitability of an indicator not on how well it represents the phenomena of interest but on how it will be evaluated by those receiving reports of this indicator. As industrial engineers, we must know when to separate measurement from evaluation. Figure 5.1 illustrates the measurement and evaluation process as having six phases, where phases 1 through 5 are measurement focused and phase 6 is evaluation. These phases are described in the following excerpt from Coleman and Clark (2001).

Phase 1—The process begins by asking what should be measured. Management or other stakeholders are interested in some event, occurrence, or phenomenon. This interest

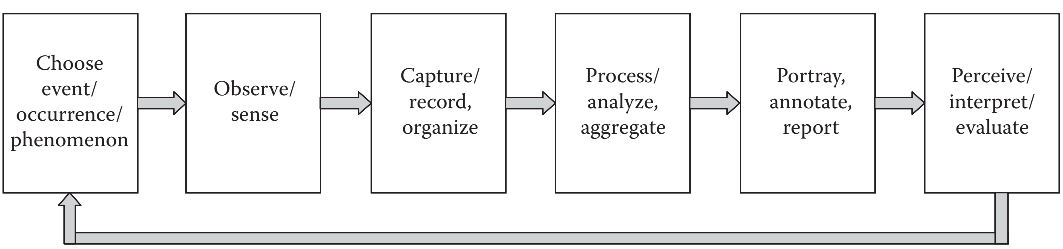

Figure 5.1 Measurement and evaluation process. (Adapted from Coleman, G.D. and Clark, L.A., A framework for auditing and assessing non-financial performance measurement systems, in *Proceedings of the Industrial Engineering Research Conference*, Dallas, CD-ROM, 2001.)

may be driven by a need to check conformity, track improvement, develop expectations for planning, diagnose problems, or promote accomplishments. This phenomenon of interest is often described in terms of key performance areas (KPAs) or criteria, which represent the priorities associated with this phenomenon.

Phase 2—The phenomenon of interest is observed or sensed to measure each KPA. One or more indicators may be measured to represent the KPA. Each indicator requires an operational definition (a defined procedure for how the observation will be converted into data). While the KPAs are "glittering generalities," the indicators are specific and reliable.

Phase 3—The output of the measurement procedure is data, which are then captured or recorded for further use. Capturing represents entering the data into the "system," whether a paper or an electronic system. This step includes ensuring that all the data generated are captured in a timely, consistent, and accurate manner. This often includes organizing or sorting the data (by time, place, person, product, etc.) to feed the analysis procedures.

Phase 4—Raw data are analyzed or processed to produce information. Manual calculations, spreadsheets, statistical software packages, and other tools are used to summarize and add value to the data. Summarizing often includes aggregating data across time or units. That is, individual values are captured and processed; then, totals or means are calculated for reporting.

Phase 5—The output of analyzing the data is information, portrayed in the format preferred by the user (manager). That is, when the values of the indicators representing KPAs for a particular phenomenon are measured, the portrayal should provide context that helps the user understand the information (Wheeler, 1993). Too often, the analyst chooses a portrayal reflecting his or her own preference rather than the user's preference.

Phase 6—The last step of the measurement and evaluation process is to perceive and interpret the information. How the user perceives the information is often as much a function of portrayal as content (see Tufte's [1997a,b] work for outstanding examples of the importance of portrayal). Regardless of which requirement (checking, improvement, planning, diagnosis, or promotion) prompted measurement, it is the user's perception of the portrayed information that is used to evaluate the performance of the phenomenon of interest. Evaluation results in continued measurement and evaluation, redesign of how the phenomenon is measured, or discontinuation and perhaps a transfer of interest to another phenomenon (Coleman and Clark, 2001).

5.4 Purposes of strategic performance measurement

Effective measurement demands that everyone understand why the measurement system is being created and what is expected from it. Design questions that arise during measurement system development can often be answered by referring back to the purpose of the system. Equally important is identification of all the users of the measurement system. If the system is being created for control purposes, then the manager or management team exerting control is the primary user. If the system is being created to support improvement, then most of or the entire unit being measured may be users. The users should be asked how they will use the measurement system. Specifically, what kinds of decisions do they intend to make on the basis of the information they receive from the measurement system? What information (available now or not) do they feel they need to support those decisions? The effectiveness of performance measurement is often dependent on how well its purpose and its user set are defined. That is, when one is evaluating whether a particular indicator is a "good" performance measure, one must first ask who will use it and what the intended purpose of the indicator is. An indicator that is good for one purpose or one user may not be as effective for another. Alternatively, an indicator that is potentially good for two or more purposes may best be used for only one purpose at a time. The use of the same indicator for potentially competing purposes, even though it could meet either purpose under ideal conditions, may lead to distortion (tampering), reluctance to report performance, or unexpected consequences, such as a lack of cooperation among the units being measured. In organizations, performance is typically measured for one or more of the following purposes:

- Control
- Improvement
- Planning
- Diagnosis
- Promotion

5.4.1 Control

Measuring performance for control may be viewed as measuring to check that what is expected has in fact occurred. Typically, a manager uses control indicators to evaluate the performance of some part of the organization for which the manager is responsible, such as a plant or department. A higher-level manager may have multiple units to control and require separate indicators from each unit. A lower-level manager may use indicators to control the performance of the individuals who work directly for that manager. In either case, the individual or unit whose performance is being monitored and controlled reports performance "upline" to the manager. If another part of the organization has the measurement responsibility (e.g., accounting and finance, quality control, or internal audit), it reports the most recent value of the indicators to the manager. The manager then reviews the level of performance on these indicators to check if the expectations are being met. Depending on the results of the comparison of current performance to expectations, and the manager's personal preferences, the manager takes action (or not) to intervene with the unit for the purpose of changing future levels of performance. Too often, managers only provide substantial feedback to the unit being evaluated when performance does not meet expectations. Control can be better maintained and performance improved when managers also reinforce good performance by providing feedback on expectations that are being met.

Care should be taken to distinguish between using an indicator to control the performance of an organizational unit and using the same indicator to judge the performance of the individuals managing or working in that unit. Measures of performance needed by managers may include elements of performance not completely within the control of those managing and working in that unit. For example, an indicator of total revenue generated by a plant may reflect the effectiveness of ongoing sales efforts, unit pricing pressure in the market, or a temporary downturn in the economy. While taking action in response to any of these factors may be appropriate for the senior-level manager who checks this plant's performance, judging the performance of local managers at the plant level by total revenue could lead to an emphasis on "making the numbers" over focusing on the factors that the local managers do control. "Making the numbers" in this situation could lead to such potentially undesirable consequences as building to inventory or spending for overtime to meet increased production targets generated by lower sales prices. A good rule of thumb is to measure performance one level above the level of control over results to encourage strategic action and to avoid suboptimization. At the same time, judgment of the performance of individual managers should focus on the causes and effects they control within the context of overall organizational performance. It is the leadership's job to assist these managers in dealing with the factors beyond their control that affect their unit's overall performance.

5.4.2 Improvement

Measuring performance for improvement is more internally focused than measuring for control. Measuring for improvement focuses on measuring the performance of the unit one is responsible for and obtaining information to establish current performance levels and trends. The emphasis here is less on evaluating something or someone's performance, and more on understanding current performance levels, how performance is changing over time, the impact of managerial actions, and identifying opportunities for improving performance. Managers often measure a number of things for use by themselves and their subordinates. An astute manager will identify drivers of end-result performance (e.g., sales, profits, customer warranty claims) and develop indicators that lead or predict eventual changes in these end results. Such leading indicators might include employee attitudes, customer satisfaction with service, compliance with quality management systems, and percent product reworked. Sears found that changes in store-level financial results could be predicted by measuring improvements in employee attitudes toward their job and toward the company. This predicted employee behavior, which, in turn, influenced improvements in customer behavior (customer retention and referral to other customers), leading, finally, to increases in revenue and operating margin (Rucci et al., 1998).

Employees, supervisors, and managers should be encouraged to establish and maintain indicators that they can use as yardsticks to understand and improve the performance of their units, regardless of whether these indicators are needed for reporting upline. Simply measuring a key performance indicator and making it promptly visible for those who deliver this performance can lead to improvement with little additional action from management. This assumes that those who deliver this performance know the desired direction for improvement on this indicator and have the resources and discretion to take actions for improvement. It is the leadership's job to make sure the people in the organization have the knowledge, resources, discretion, and direction to use performance information to make improvements.

5.4.3 Planning

Measuring for the purpose of planning has at least two functions: (1) increasing understanding of current capabilities and the setting of realistic targets (i.e., goals) for future performance; and (2) monitoring progress toward meeting existing plans. One could argue that these simply represent planning-centric versions of measuring for improvement and then measuring for control. The role of measuring performance as part of a planning effort is important enough to warrant a separate discussion.

Nearly all strategic management or strategic planning efforts begin with understanding the organization and its environment. This effort is referred to as internal and external strategic analysis (Thompson and Strickland, 2003), organizational systems analysis (Sink and Tuttle, 1989), or, in plain words, "preparing to plan." A key part of internal analysis is understanding current performance levels, including the current value of key performance indicators and their recent trends. This provides the baseline for future performance evaluations of the effectiveness of the planned strategy and its deployment. Also, the choice of key performance indicators tells the organization what is important and is a specific form of direction often more carefully followed than narrative statements of goals and vision. Understanding current performance and its relation to current processes and resources provides managers with a realistic view of what is possible without having to make substantial changes to the system. Thus, setting intelligent targets for future performance requires an understanding of how implementation of the plan will change processes and resources to enable achievement of these targets. A key part of the external analysis is obtaining relevant comparisons so that the competitiveness of current performance levels and future performance targets can be evaluated. To answer the question of how good a particular performance level is, one must ask "compared to what?" Current competitor performance provides an answer to this question, but it must be assumed that competitors are also planning for improved performance. Setting future performance targets must take this moving competitive benchmark into account. Even the projected performance of your best current competitor may be inadequate as a future performance target to beat. The strategic management literature is full of examples of corporations that did not see their new competition coming and were blindsided by new competitors playing by different rules with substitutes for these corporations' bread-and-butter products (see Hamel and Prahalad, 1996; Hamel, 2002). As Drucker (1998) has pointed out, some of the most important information managers need comes from outside their organizations and even outside their industries. A challenge for performance measurement is to provide not only internal but also external performance information that provides competitive intelligence for making strategic decisions.

Most strategic management or strategic planning processes include a last or next to last step that serves to measure, evaluate, and take corrective action. Often, this step is expected to be occurring throughout the process, with the formal execution of the explicit step occurring after goals have been set, action plans deployed, and strategy implementation is under way. That is, periodic review of progress toward meeting goals is a regular part of a strategic management effort, and performance indicators can provide evidence of that progress. When the goal setting process includes the identification of key performance indicators and future performance targets for each indicator, the decision of which indicators to review has largely been made. In cases where goals are perhaps more qualitative or include simple quantitative targets without an operationally defined performance indicator, the planning team must choose or develop a set of progress indicators for these periodic (e.g., monthly or quarterly) reviews. A rule of thumb for these cases, based on the work of Sink and Tuttle (1989), is to develop indicators that provide evidence of the effectiveness, efficiency, quality, and impact of progress on each goal. Each of these terms is defined earlier. Even when key performance indicators have been predetermined at the time of goal setting, additional "drill-down" indicators may be required to explain performance trends and illustrate perceived cause-and-effect relationships among managerial actions, environmental and competitor actions, and observed levels of performance on end-result indicators.

Once the indicators have been chosen or developed, the periodic reviews are much more than collecting data, reporting current performance levels, and comparing to plan. How these reviews are conducted has a major impact on the organization's approach and even success with strategic management. If the reviews overemphasize checking or making sure that the people responsible for each goal are making their numbers, then reviews run the risk of taking on a confrontational style and may lead to gaming, distortion, and hoarding of information. On the other hand, reviews that focus on what can be learned from the performance information and sharing lessons, and even resources when needed, can lead to better goal setting, improved action plans for implementing strategies, and increased sharing of performance information that may indicate future trends, good or bad. The type of review chosen is likely to reflect the organization's culture and the leadership's preferences. While either style may be used to drive performance, the two styles differ in the types of initiatives and actions leadership must take outside of and between periodic reviews to support performance improvement.

5.4.4 Diagnosis

Measuring performance for diagnosis or screening (Thor, 1998) is similar to the drill-down described for illustrating cause-and-effect relationships among controllable and non-controllable factors and their impact on end results. When an undesired (or desirable but unexplainable) result on a key indicator is observed, exploring the recent history of related indicators may provide insight into the possible causes. Tools such as the cause-and-effect (fishbone) diagram (Brassard and Ritter, 1985; Ishikawa, 1985) or quality function deployment (Akao, 1990) are useful in identifying drill-down metrics, likely to be at the cause of the observed effect. Unlike the previous methods, which are used for continual measurement of performance, measuring for diagnosis may be a one-time measurement activity with a start and an end. Thus, devoting resources to systematizing or institutionalizing the new indicators required should be based on the likelihood that these indicators will be needed again in the near future. When assessing the indicators of an existing measurement system, look for indicators once needed for diagnosis that have outlived their usefulness; stopping those outdated indicators may free up resources needed to produce newly identified indicators.

5.4.5 Promotion

Measuring for promotion (an idea contributed by Joanne Alberto) is using performance indicators and historical data to illustrate the capabilities of an organization. The intent is to go beyond simple sales pitch claims of cutting costs by *X*% or producing product twice as fast as the leading competitor. Here, the manager is using verifiable performance information to show the quantity and quality of product or service the organization is capable of delivering. Not only does this performance information show what is currently possible, it also provides a potential client with evidence that the organization measures (and improves) its performance as part of its management process. Thus, the customer can worry less about having to continually check this provider's performance and can rely on the provider to manage its day-to-day performance. A caveat here is that it is important to balance the organization's need to protect proprietary performance information with the customer's need for evidence of competitive product and service delivery. Care should also be taken in supporting the validity of promotional performance information so that the claims of less scrupulous competitors, who may boast of better levels of performance but present poorly substantiated evidence, are discounted appropriately.

Once the manager or engineer has clarified why performance is being measured, the question of what to measure should be addressed. Organizational performance is multi-dimensional, and a single indicator rarely meets all the needs of the intended purpose.

5.5 Dimensions of performance

This section describes a number of frameworks for organizing the multiple dimensions of organizational performance. Each framework is a useful tool for auditing an organization's collective set of indicators to identify potential gaps. The intent here is neither to advocate the adoption of a specific framework as the measurement categories for a given organization, nor to advocate that an organization has at least one indicator for every dimension of these frameworks. The astute management team must recognize that organizational performance is multidimensional and make sure their measurement system provides performance information on the dimensions key to the success of their organization.

Those interested in a philosophical discussion of performance dimensions and how to choose the appropriate unit of analysis should read Kizilos' (1984) "Kratylus automates his urn works." This thought-provoking article sometimes frustrates engineers and managers who are looking for a single "correct" answer to the question of what dimensions of performance should be measured. The article is written in the form of a play with only four characters and makes excellent material for a group discussion or exercise.

5.5.1 Concept of key performance areas

Key performance areas are the vital few categories or dimensions of performance for a specific organization. KPAs may or may not reflect a comprehensive view of performance, but they do represent those dimensions most critical to that organization's success. While the indicators used to report the performance of each KPA might change as strategy or the competitive environment changes, the KPAs are relatively constant.

Rather than simply adopting one of the performance dimension frameworks described in this section, an organization's managers should familiarize themselves with the alternative frameworks and customize the dimensions of their organizational scoreboard to reflect their organization's KPAs. What is most important is that the measurement system provides the managers with the information necessary to evaluate the organization's performance in all key areas (i.e., KPAs) as opposed to conforming to someone else's definition of balance.

5.5.2 Balanced scorecard

While it has long been recognized that organizational performance is multidimensional, the practice of measuring multiple performance dimensions was popularized by the introduction of Kaplan and Norton's (1992) Balanced Scorecard. At its core, the Balanced Scorecard recognizes that organizations cannot be effectively managed with financial measures alone. While necessary for survival, financial measures tend to be lagging indicators of results and are frequently difficult to link to managerial actions aimed at improving medium- to long-term performance. Compounding this shortcoming, financial measurement systems are typically designed to meet reporting requirements for publicly traded companies or auditor's requirements for government agencies and privately held companies (i.e., financial accounting). Providing information to support managing the organization (i.e., managerial accounting) is an afterthought. This creates a situation where indicators developed for one purpose (fiscal control) are reused for another purpose (management and improvement), creating predictable problems.

The Balanced Scorecard views organizational performance from four perspectives, with the financial perspective being one of those four. The other three perspectives are the customer perspective, the internal process perspective, and the learning and growth perspective. Kaplan and Norton (1996) later suggested a general causal structure among the four perspectives. Thus, managerial actions to improve learning and growth, both at the individual and organizational level, should result in improved performance on indicators of internal process performance, assuming the learning and growth initiatives and indicators are aligned with the internal process objectives. Improved performance on internal process indicators should result in improved results of the customer perspective indicators, if the process indicators reflect performance that is ultimately important to customers. And finally, if the customer perspective indicators reflect customer behaviors likely to affect the organization, then it is reasonable to expect improved performance on these customer indicators to lead to improved financial performance. For example, an initiative aimed at improving the quality assurance skills of quality technicians and quality management skills of production line supervisors might be indicated by increased numbers of Certified Quality Technicians and Certified Quality Managers (learning and growth indicators). Assuming this initiative was aimed at closing a relevant gap in skills, the application of these skills could be expected to improve levels of internal process indicators such as percent scrap and shift the discovery of defects further upline in the value stream (potentially reducing average cycle time for good product produced). Improvements in results on these internal process indicators could lead to fewer customer warranty returns, translating into direct financial savings. Improved performance on other customer-perspective indicators such as customer perceptions of quality and their likelihood to recommend the product to others, although less directly linked, may also be predictors of improved financial results such as increased sales.

While popular, the Balanced Scorecard has received some valid criticism. Nørreklit (2003) argues that the Balanced Scorecard has generated attention on the basis of persuasive rhetoric rather than on convincing theory. Theoretical shortcomings include suggested cause-and-effect relationships based on logic rather than empirical evidence and use of a strategic management system without addressing key contextual elements of strategic management (e.g., monitoring key aspects of the dynamic external environment or employing a top–down control model for implementation that appears to ignore organizational realities). Pfeffer and Sutton (2000, p. 148) point out that the Balanced Scorecard is "great in theory," but identify a number of problems in its implementation and use: "The system is too complex, with too many measures; the system is often highly subjective in its actual implementation; and precise metrics often miss important elements of performance that are more difficult to quantify but that may be critical to organizational success over the long term."

The industrial engineer's challenge is to sort through these shortcomings and address them with a well-designed measurement system that aligns with other management systems and balances practical managerial needs with theoretical purity. Practical issues related to designing and implementing a measurement system are described previously.

Richard Barrett (1999a, 1999b) proposed enhancing the Balanced Scorecard by expanding the customer perspective to include suppliers' perspectives and adding three additional perspectives: corporate culture, community contribution, and society contribution. Certainly the importance of supply chain management and partnering with suppliers warrants the inclusion of a supplier perspective in an organizational scorecard. Corporate culture has long been recognized as important to organizational success (Deal and Kennedy, 1982; Peters and Waterman, 1982) and appears as a key factor in the popular press accounts of great organizations. However, much work remains regarding how best to measure corporate culture and to use this information to better manage the organization. Management scholar Ralph Kilmann (1989; Kilmann and Saxton, 1983) and industrial engineer Larry Mallak (Mallak et al., 1997; Mallak and Kurstedt, 1996) offer approaches to measuring corporate culture. Off-the-shelf survey instruments, such as Human Synergistics® International's Organizational Culture Inventory®, are also available. Barrett's recommendation to measure community and societal contributions are similar dimensions measured at different levels. Community contribution includes not only the cities, counties, and states where the organization and its employees reside and sell their products but also the industries and professions in which the organization operates. Societal contribution expands beyond local impact and measures the organization's immediate and longer-term global impact.

The industrial engineer should recognize that the Balanced Scorecard is only one framework for organizing the dimensions of organizational performance, and should be familiar with various alternatives and develop or adapt a framework that fits the organization's needs.

5.5.3 Baldrige criteria

A widely accepted performance dimensions framework that is updated bi-annually is the Results category of the Baldrige Criteria for Performance Excellence (Baldrige Performance Excellence Program, 2013). This framework consists of five items that may be thought of as performance dimensions: product and process results, customer-focused results, workforce-focused results, leadership and governance results, and financial and market results. When identifying indicators for each dimension, the Baldrige criteria stress choosing indicators that are linked to organizational priorities such as strategic objectives and key customer requirements. The criteria emphasize segmenting results to support meaningful analysis and providing comparative data to facilitate the evaluation of levels and trends. The Baldrige criteria also include relative weights for each of these dimensions.

Indicators of product and process results provide evidence of the performance of products and processes important to customers. In the food service industry where customers want healthy eating alternatives, this might include providing comparisons of nutritional information of your leading products to those of key competitors. Process results also include process effectiveness results for strategy and operations. Indicators of customer-focused results provide evidence of the attitudes and behaviors of customers toward a company's products and services. This requires not only indicators of customer satisfaction and dissatisfaction but also indicators of customer engagement such as their willingness to recommend the company's products to others. Workforce-focused results are particularly relevant to industrial engineers because they include indicators of workforce capability and capacity, and workforce engagement. Industrial engineers address the organization and management of work, including how work and jobs are organized and managed to create and maintain "a productive, caring, engaging, and learning environment for all members of your workforce" (Baldrige Performance Excellence Program, 2011, p. 48). Measuring the levels and trends of workforce capability and capacity could be an indicator of the performance of the industrial engineering function. Other items to be reported under workforce-focused results include indicators of workforce climate such as safety and absenteeism, workforce engagement such as turnover and satisfaction, and workforce and leader development such as number of certifications and promotions. Such indicators are not just the domain of the human resource manager, but include indicators that reflect the effectiveness of the work systems and supporting aids developed by the industrial engineers. The leadership and governance results dimension starts with indicators of leadership communication and engagement to deploy vision and values and create a focus on action. Indicators providing evidence of effective governance and fiscal accountability might include financial statement issues and risks, and important auditor findings. This dimension also includes social responsibility results, addressing evidence of achieving and passing regulatory and legal requirements, indicators of ethical behavior and stakeholder trust, and indicators of the organization's support of its key communities. The final dimension in the Baldrige results framework is financial and market results. This dimension includes traditional financial indicators such as return on investment and profitability, and market indicators such as market share and market share growth.

5.5.4 Sink's seven criteria

D. Scott Sink provides an industrial engineer's view of performance with his seven performance criteria (Sink, 1985; Sink and Tuttle, 1989). He suggests that organizational performance can be described in terms of seven interrelated criteria:

- Effectiveness: indicators of doing the correct things; a comparison of actual to planned outputs
- Efficiency: a resource-oriented criterion; a comparison of planned to actual resources used
- Quality: defined by one or more of David Garvin's (1984) five definitions of quality (transcendent, product-based, manufacturing-based, user-based, or value-based) and measured at up to five (or six) points throughout the value stream
- Productivity: an indicator based on a ratio of outputs to the inputs required to produce those outputs (more on productivity later)
- Innovation: indicators of organizational learning and growth as applied to the organization's current or future product and service offerings
- Quality of work life: indicators of employee-centered results; preferably those predictive of higher levels of employee work performance
- Profitability/budgetability: indicators of the relationship of revenues to expenses; whether the goal is to make a net profit or to stay within budget (while delivering expected levels of service)

5.5.4.1 Productivity

Productivity is a particularly important concept for industrial engineers and warrants further discussion here. Productivity indicators reflect the ratio of an organization's or individual's outputs to the inputs required to produce those outputs. The challenge is determining which outputs and inputs to include and how to consolidate them into a single numerator and denominator. Outputs include all the products and services an organization produces and may even include by-products. Inputs include labor, capital, materials, energy, and information.

Many commonly used productivity indicators are actually partial measures of productivity. That is, only part of the total inputs used to produce the outputs are included in the denominator. The most common are measures of labor productivity, where the indicator is a ratio of outputs produced to the labor inputs used to produce them (e.g., tons of coal per man day, pieces of mail handled per hour). While relatively simple and seemingly useful, care should be taken in interpreting and evaluating the results of partial productivity indicators. The concept of input substitution, such as increasing the use of capital (e.g., new equipment) or materials (e.g., buying finished components rather than raw materials), may cause labor productivity values to increase dramatically, owing to reasons other than more productive labor. A more recent shortcoming of measuring labor productivity is that direct labor has been steadily decreasing as a percent of total costs of many manufactured, mined, or grown products. In some cases, direct labor productivity today is at levels almost unimaginable 20 or 30 years ago. One might argue that the decades-long emphasis on measuring and managing labor productivity has succeeded, and that industrial engineers in these industries need to turn their attention to improving the productivity of materials and energy, and perhaps indirect labor. For more information, Sumanth (1998) provides a thoughtful summary of the limitations of partial productivity measures.

Total or multifactor productivity measurement approaches strive to address the limitations of partial productivity measures. Differing outputs are combined using a common scale such as constant value dollars to produce a single numerator, and a similar approach is used to combine inputs to produce a single denominator. Total factor approaches include all identifiable inputs, while multifactor approaches include two or more inputs, typically the inputs that make up the vast majority of total costs. The resulting ratio is compared with a baseline value to determine the percent change in productivity. Miller (1984) provides a relatively simple example using data available from most accounting systems to calculate the changes in profits due to any changes in productivity, as well as to separate out profit changes due to price recovery (i.e., net changes in selling prices of outputs relative to the changes in purchasing costs of inputs). Sink (1985) and Pineda (1996) describe multifactor models with additional analytical capabilities, useful for setting productivity targets based on budget targets and determining the relative contributions of specific inputs to any changes in overall productivity. Other approaches to productivity measurement such as data envelopment analysis (DEA) are beyond the scope of this chapter. See Cooper et al. (2004) and Medina-Borja et al. (2006) for further information about the use of DEA.

5.5.4.2 *Quality*

Quality, like productivity, deserves additional attention in an industrial engineer's view of measuring performance. Quality is ultimately determined by the end user of the product or service. And often, there are many intermediate customers who will judge and perhaps influence the quality of the product before it reaches the end user. As there are numerous definitions of quality, it is important to know which definition your customers are using. While your first customer downstream (e.g., an original equipment manufacturer or a distributor) might use a manufacturing-based (i.e., conformance to requirements) indicator such as measuring physical dimensions to confirm they fall within a specified range, the end user may use a user-based (i.e., fitness-for-use) indicator such as reliability (e.g., measuring mean time between failures) to evaluate quality. A full discussion of the five common definitions of quality and the eight dimensions of quality (performance, features, reliability, conformance, durability, serviceability, aesthetics, and perceived quality) is found in Garvin (1984). While seemingly adding confusion to the definition of quality within a larger performance construct, Garvin's eight dimensions of quality can be thought of as differing perspectives from which quality is viewed. Without multiple perspectives, one may get an incomplete view of a product's quality. As Garvin points out, "a product can be ranked high on one dimension while being low on another" (p. 30).

Once one or more definitions of quality have been chosen, the industrial engineer must decide where to measure quality before finalizing the indicators to be used. Sink and Tuttle (1989) describe quality as being measured and managed at five (later six) checkpoints. The five checkpoints correspond to key milestones in the value stream, with checkpoints 2 and 4 representing traditional incoming quality measurement (before or just as inputs enter the organization) and outgoing quality measurement (just before outputs leave the organization), respectively. Quality checkpoint 3 is in-process quality measurement, a near-discipline in its own right, including statistical process control methods, metrology, certified quality technicians, and certified quality engineers. At checkpoint 3, we are measuring the key variables and attributes of processes, products, and services that predict or directly lead to the desired characteristics at outgoing quality measurement (quality checkpoint 4) as well as those that contribute to success on the quality dimensions that are important further downstream (see checkpoint 5). Tracking such variables and attributes lends itself to statistical analysis. See Chapters 3 and 36 for a discussion of statistical process control. For an excellent introduction to applying statistical thinking and basic methods to management data, see Donald Wheeler's *Understanding Variation* (1993). The novice industrial engineer can benefit by taking heed of the late W. Edwards Deming's (1993) often stated admonition to begin by "plotting points" and utilizing the "most under-used tools" in management, a pencil and piece of grid paper. Quality checkpoint 1 is proactive management of suppliers and includes the indicators used to manage the supply chain. What might be incoming, in-process, outgoing, or overall quality management system indicators from the supplier's perspectives are quality checkpoint 1 indicators from the receiving organization's perspective. Quality checkpoint 5 is the measurement of product and service quality after it has left the organization's direct control and is in the hands of the customers. Quality checkpoint 5 might include indicators from the Baldrige items of product and service outcomes and customer-oriented results. Quality checkpoint 5 indicators provide evidence that products or services are achieving the outcomes desired by customers and the customer's reactions to those outcomes. The sixth, sometimes omitted, checkpoint is measuring the overall quality management or quality assurance process of the organization. Today we may relate this sixth checkpoint to the registration of an organization's quality management systems, as evidenced by receiving an ISO 9001 certificate.

5.5.5 Human capital

Industrial engineers have long been involved in the measurement and evaluation of the performance of individuals and groups. As the knowledge content of work has increased, the overall cost and value of knowledge workers has increased. Organizations spend substantial energy and resources to hire, grow, and retain skilled and knowledgeable employees. Although these expenditures are likely to appear in the income statement as operating costs, they are arguably investments that generate human capital. While an organization does not own human capital, the collective knowledge, skills, and abilities of its employees do represent an organizational asset—one that should be maintained or it can quickly lose value. Organizations need better measurement approaches and performance indicators to judge the relative value of alternative investments that can be made in human capital. They need to know which are the most effective options for hiring, growing, and keeping talent. The following paragraphs provide the industrial engineer with context and examples to help tailor their performance measurement toolkit to the unique challenges associated with measuring the return on investments in human capital.

Traditional human resource approaches to measuring human capital have focused on operational indicators of the performance of the human resources function. In particular, these indicators have emphasized the input or cost side of developing human capital. Such indicators might include average cost to hire, number of days to fill an empty position, or cost of particular employee benefits programs. More holistic approaches (Becker et al., 2001) focus on business results first, and then link indicators of how well human capital is being managed to create those results.

Assuming the organization has developed a multidimensional performance measurement system as described in this chapter, the next step is to identify human capital-related drivers of the leading organizational performance indicators (e.g., product and process outcomes, customer-focused outcomes as opposed to lagging performance results such as financial and market outcomes). Such drivers are likely to be related to employee attitudes and behaviors. Drivers of customer-focused outcomes might include employee attitudes toward their jobs or supervisors, or behaviors such as use of standard protocols and knowing when to escalate an issue to a customer service manager. Drivers of product and process outcomes might include behaviors such as use of prescribed quality assurance procedures, completing customer orientation upon delivery, or perhaps an organizational effectiveness indicator such as cycle time (i.e., where cycle time is heavily dependent on employee performance). Indicators of the health of an organization's human capital are likely to predict or at least lead performance on these human capital-related drivers of organizational performance. Indicators of the health of human capital reflect the value of human capital as an organizational asset. Examples of such indicators include average years of education among knowledge workers (assumes a relatively large pool of employees), a depth chart for key competencies (i.e., how many employees are fully qualified to fulfill each mission), attrition rates, or more sophisticated turnover curves that plot turnover rates in key positions by years of seniority. Finally, traditional cost-oriented measures of human resource programs can be evaluated in terms of their impact on the health of human capital and human capital drivers of organizational performance.

Human capital indicators should help answer questions such as the following: Does the new benefit program reduce turnover among engineers with 10–20 years of experience? Does the latest training initiative expand our depth chart in areas that were previously thin, thus reducing our risk of not being able to meet product and service commitments? Do changes to our performance management system improve employee attitudes among key customer interface employees? Do our initiatives aimed at improving employee attitudes and behaviors translate into better products and services as well as customers who increase the percentage of their business they give to our organization? Measuring human capital and the return on investments in human capital are new frontiers in measurement for industrial engineers, with the potential to make substantial contributions to organizational competitiveness.

5.6 *Implementing a measurement system*

Once clear about why to measure performance and what dimensions of performance to measure, the question becomes how to implement a functioning measurement system. The measurement system includes not only the specific indicators but also the plan and procedures for data gathering, data entry, data storage, data analysis, and information portrayal, reporting, and reviewing. A key recommendation is that those whose performance is being measured should have some involvement in developing the measurement system. The approaches that can be used to develop the measurement system include the following: (1) have internal or external experts develop it in consultation with those who will use the system; (2) have the management develop it for themselves and delegate implementation; (3) have the units being measured develop their own measurement systems and seek management's approval; or (4) use a collaborative approach involving the managers, the unit being measured, and subject matter expert assistance. This last approach can be accomplished by forming a small team, the measurement system design team.

A "design team" is a team whose task is to design and perhaps develop the measurement system; however, day-to-day operation of the measurement system should be assigned to a function or individual whose regular duties include measurement and reporting (i.e., it should be an obvious fit with their job and be seen as job enrichment rather than an add-on duty unrelated to their regular work). When ongoing performance measurement is assigned as an extra duty, it tends to lose focus and energy over time and falls into a state of neglect. Depending on how work responsibility is broken down in an organization, it may make sense to assign responsibility for measurement system operation to industrial engineering, accounting and finance, the chief information officer, quality management/assurance, human resources, or a combination of these. The design team should include the manager who "owns" the measurement system, a measurement expert (e.g., the industrial engineer), two or more employees representing the unit whose performance is being measured, and representatives from supporting functions such as accounting and information systems.

Each of the four development approaches can benefit from adopting a systems view of the organization using an input/output analysis.

5.6.1 Input/output analysis with SIPOC model

A tool for helping users identify information needs at the organizational level is the input/output analysis or the SIPOC (suppliers, inputs, processes, outputs, and customers) model. The intent is to get the users to describe their organization as an open system, recognizing that in reality there are many feedback loops within this system that make it at least a partially closed-loop system. The SIPOC model is particularly useful for the design team approach to developing a measurement system. The model helps the team members gain a common understanding of the organization and provides a framework for discussing the role and appropriateness of candidate indicators.

The first step to complete the SIPOC model is to identify the organization's primary customers, where a customer is anyone who receives a product or service (including information) from the organization. Next identify the outputs, or specific products and services, provided to these customers. For an organization with a limited number of products and services, these outputs can be identified on a customer-by-customer basis; for an organization with many products and services, it is more efficient to identify the products and services as a single comprehensive list and then audit this list customer by customer to make sure all relevant products and services are included.

The next step is not typically seen in the SIPOC model, but it is a critical part of any input/output analysis. It starts with the identification of the customers' desired outcomes, that is, the results they want as a consequence of receiving the organization's products and services. A customer who purchases a car may want years of reliable transportation, a high resale value, and styling that endures changes in vogue. A customer who purchases support services may want low-cost operations, seamless interfaces with its end users, and a positive impact on its local community. While the organization may not have full control in helping its customers achieve these desired outcomes, it should consider (i.e., measure) how its performance contributes to or influences the achievement of these outcomes. The identification of desired outcomes also includes identifying the desired outcomes of the organization, such as financial performance (e.g., target return on investment, market share), employee retention and growth, repeat customers, and social responsibility. Measuring and comparing the customer's desired outcomes to the organization's desired outcomes often highlights key management challenges, such as balancing the customer's desire for low prices with the organization's financial return targets. Measuring outcomes helps the organization understand customer needs beyond simply ensuring that outputs meet explicit specifications.

At the heart of the SIPOC model is the identification of processes, particularly the processes that produce the products and services. A separate list of support processes, those that provide internal services necessary to the functioning of the organization but are not directly involved in producing products or services for external consumption, should also be identified. Processes lend themselves to further analysis through common industrial engineering tools such as process flow charts and value stream maps. Process flow charts are useful for identifying key measurement points in the flow of information and materials and thus the source of many operational performance indicators. Strategic performance measurement may include a few key process indicators, particularly those that predict the successful delivery of products and services. Once processes are identified, the inputs required for those processes are identified. As with outputs, it may be more efficient to identify inputs as a single list and then compare them to the processes to make sure all key inputs have been identified. The five generic categories of inputs that may be used to organize the list are labor, materials, capital, energy, and information. In order to be useful for identifying performance indicators, the inputs must be more specific than the five categories. For example, labor might include direct hourly labor, engineering labor, contracted labor, management, and indirect labor. These can be classified further if there is a need to measure and manage labor at a finer level, although this seems more operational than strategic. Examples of relevant labor indicators include burdened cost, hours, percent of total cost, and absenteeism. The last component of the SIPOC model is the identification of suppliers. While this component has always been important, the advent of overt improvement approaches such as supply chain management and the increased reliance on outsourcing have made the selection and management of suppliers a key success factor for many organizations. Suppliers can also be viewed as a set of upstream processes that can be flow charted and measured like the organization's own processes. The design team may wish to work with key suppliers to identify indicators of supplier performance that predict the success of (i.e., assure) the products and services being provided as inputs in meeting the needs of the organization's processes and subsequent products and services.

Informed by the insight of working through an input/output analysis, and regardless of whether a design team is used or not, the process of designing, developing, and implementing a measurement system based on the body of knowledge described thus far is conceptually simple and practically quite complex. An outline of the sequential steps in this process is provided as a guide in the following section.

5.6.2 Macro strategic measurement method

There are essentially seven steps in the process of building and using a strategic measurement system. Each of these seven macro steps may be decomposed into dozens of smaller activities depending on the nature and characteristics of the organization. In practice, the steps and substeps are often taken out of sequence and may be recursive.

1. Bound the target system for which performance measures will be developed. This seemingly obvious step is included as a declaration of the importance of operationally and transparently defining the system of interest. Is the target system a single division or the entire firm? Are customers and suppliers included in the organizational system or not? Are upline policy makers who influence the environment inside the system or outside of it? Any particular answer may be the "right" one; the important point is shared clarity and agreement. Frequently people who want better measurement systems define the target system too small, in the false belief

that it is inappropriate to measure things that may be out of the target system's control. The false belief is often present at the functional and product level, and at the organizational level as supply chains become more complex. Indicators that reflect performance only partially controllable or influenced by the organization are often those most important to customers and end users. When the organization has only partial control of a performance indicator of importance to customers, the organization needs to understand its contribution to that performance and how it interacts with factors beyond its control. This aversion to measure what is outside one's control is driven by an inability to separate measurement from evaluation. To separate the two, first, measure what's important; second, evaluate performance and the degree of influence or control users have over changing the measured result.

2. Understand organizational context and strategy. This step involves documenting, verifying, or refining the target system's mission, vision, values, current state, challenges, and long- and short-term aims—all of the activities associated with strategic planning and business modeling. Recall how to do measurement in the context of planning and also the input/output process presented earlier.
3. Identify the audience(s) and purpose(s) for measuring. A helpful maxim to guide development of strategic planning and strategic measurement PDSA systems is audience + purpose = design. Who are the intended audiences and users of the measurement system, and what are their needs and preferences? What are the purpose(s) of the measurement system being developed? Effective measurement system designs are derived from those answers. There are many ways to discover and articulate who (which individuals and groups) will be using the measurement system, why they want to use it, and how they want to use it. Conceptually, the fundamental engineering design process is applicable here, as are the principles of quality function deployment for converting user needs and wishes into measurement system specifications and characteristics.
4. Select KPAs. This step involves structured, participative, generative dialogue among a group of people who collectively possess at least a minimally spanning set of knowledge about the entire target system. The output of the step is a list of perhaps seven plus or minus two answers to the following question: "In what categories of results must the target system perform well, in order to be successful in achieving its aims?"
5. For each KPA, select key performance indicators (KPIs). This step answers the question for each KPA, "What specific quantitative or qualitative indicators should be tracked over time to inform users how well the target system is performing on this KPA?" Typically a candidate set of indicators is identified for each KPA. Then a group works to clarify the operational definition and purpose of each candidate KPI; evaluate proposed KPIs for final wording, importance, data availability, data quality, and overall feasibility; consider which KPIs will give a complete picture while still being a manageable number to track (the final "family of measures" will include at least one KPI for each KPA); select final KPIs that will be tracked; and identify the KPI "owner," sources of data, methods and frequency of reporting, and reporting format for selected KPIs. An inventory of existing performance indicators should be completed in this step.

   A note on steps 4 and 5: The order of these steps as described implies a top–down approach. However, reversing the order into a bottom–up approach can also be successful. A bottom–up approach would identify candidate indicators, perhaps using a group technique such as brainstorming or the nominal group technique (Delbecq et al., 1975). Once there is a relatively comprehensive list of candidate indicators, the list can be consolidated using a technique such as affinity diagrams (Kubiak and Benbow, 2009) or prioritized with the nominal group technique or analytical hierarchy process. The aim here is to shorten the candidate list to a more manageable size by clustering the indicators into categories that form the foundation for the dimensions of the organization's scoreboard (i.e., KPAs) or a prioritized list from which the "vital few" indicators can be extracted and then categorized by one or more of the performance dimensions frameworks to identify gaps. In either case (top–down or bottom–up), the next step is to try the indicators out with users and obtain fitness-for-use feedback.

6. Track the KPIs on an ongoing basis. Include level, trend, and comparison data, along with time-phased targets to evaluate performance and stimulate improvement. Compare and contrast seemingly related KPIs over time to derive a more integrated picture of system performance. An important part of piloting and later institutionalizing the vital few indicators is to develop appropriate portrayal formats for each indicator. What is appropriate depends on the users' preferences, the indicator's purpose, and how results on the indicator will be evaluated. User preferences may include charts versus tables, use of color (some users are partially or fully color-blind), and the ability to drill down and easily obtain additional detail. An indicator intended for control purposes must be easily transmissible in a report format and should not be dependent on color (the chart maker often loses control of the chart once it is submitted, and color charts are often reproduced on black-and-white copiers), nor should it be dependent on verbal explanation. Such an indicator should also support the application of statistical thinking so that common causes of variation are not treated as assignable causes, with the accompanying request for action. An indicator intended for feedback and improvement of the entire organization or a large group will need to be easily understood by a diverse audience, large enough to be seen from a distance, and easily dispersed widely and quickly. Rules of thumb for portraying performance information are provided in Table 5.1. Not all of the considerations in Table 5.1 can be applied to every chart. A detailed discussion of portrayal is beyond the scope of this chapter. Design teams should support themselves with materials such as *Wheeler's Understanding Variation* (1993) and Edward Tufte's booklet, *Visual and Statistical Thinking: Displays of Evidence for Decision Making* (1997a), a quick and entertaining read on the implications of proper portrayal.
7. Conduct review sessions. A powerful approach to obtain feedback from users on the indicators, and to evaluate organizational performance based on the indicators, is to conduct regular periodic face-to-face review sessions. Review sessions are typically conducted with all the leaders of the target system participating as a group. Notionally, the review sessions address four fundamental questions: (1) Is the organization producing the results called for in the strategy? (2) If yes, what's next; and if no, why not? (3) Are people completing the initiatives agreed to when deploying the strategy? (4) If yes, what's next; if no, why not? The review session is where critical thinking and group learning can occur regarding the organizational hypothesis tests inherent in strategy. If predicted desired results are actually achieved, is it because leaders chose a sound strategy and executed it well? To what degree was luck or chance involved? If predicted results were not achieved, is it because the strategy was sound yet poorly implemented? Or well implemented but results are delayed by an unforeseen lag factor? Or, in spite of best intentions, did leaders select the "wrong"

Table 5.1 Rules of Thumb for Portraying Performance Information

- A picture is often worth a thousand words, so charts, sketches, and photographs should be used when they meet user needs.
- Start by developing the chart on paper (by hand), before moving to computer-generated graphics. Starting with computer-generated charts often leads to a portrayal based on what the tool can do rather than what the user desires.
- ALL CAPS IS HARDER TO READ AND IMPLIES SHOUTING; thus, use uppercase and lowercase text.
- An accompanying table of the data used to produce the chart is desirable whenever possible.
- Longitudinal data are always preferable. If a change in process or product results in a capability that is no longer comparable, annotate this change in capability and continue to show historical performance until the new capability is well established.
- For high-level indicators that aggregate performance or only indicate end results, driver indicators that provide an explanation of changes observed in the high-level indicator should be provided as supporting material (to support cause-and-effect thinking).
- Indicators should help the user understand the current level of performance, the trend in performance, and provide appropriate comparisons for evaluation. Comparisons with the performance of competitors, customer expectations, or targets set by the organization provide context for judging the desirability of results.
- When using labels to note acceptable ranges of variability, clearly distinguish limits based on the capability of the process from limits established by customers (i.e., specifications) and limits established by management (i.e., targets).
- The date produced or revised and the owner (producer) of the indicator should be clearly labeled.
- Supporting information such as formulae used, data sources, and tools used to process the data should be available as a footnote or hyperlink, or in supporting information such as an appendix.
- To the extent possible, keep portrayal formats consistent from reporting period to reporting period. Continuous improvement is laudable, but users spend more time interpreting results and making decisions when they are familiar with the format of the indicator.
- Annotate charts with the initiation and completion of improvement interventions intended to change the level, trend, or variability of results.
- Acknowledge possible omissions or errors in the data as part of the portrayal.

strategy? Group discussion of these strategy and measurement questions will also cause further suggestions to be made to enhance the set of indicators and how they are portrayed. See Farris et al. (2011) for more on review sessions.

5.7 *Performance measurement pitfalls*

Performance measurement may seem rational and logical, yet implementation of many performance measurement systems fails. Here are some of the pitfalls that can contribute to failure. The reader should note that many of these pitfalls are related to the motivational aspects of measuring and evaluating performance.

- A standard set of measurements created by experts will not help. A method is needed by which measurement teams can create and continually improve performance measurement systems suited to their own needs and circumstances.

- Participation in the process of designing and implementing a performance measurement system facilitates its implementation and enhances its acceptance.
- To be "built to last," the measurement system must support decision making and problem solving.
- A documented and shared definition of the target system for the performance measurement effort is essential for success, as are well-crafted operational definitions for each measure of performance.
- Visibility and line-of-sight must be created for measurement systems to ensure effective utilization.
- Measurement is often resisted. Some reasons for this resistance include the following:
  - Data are collected but not used. It is important to be mindful that the purpose of measurement is not to generate data needlessly, but to generate data that can actually inform future decision making.
  - Fear of the consequences of unfavorable results.
  - Fear of the consequences of favorable results, such as justifying a reduction in resources.
  - Leaders ask "What will we do if our results are bad?" The answer is simple: you use this information as an opportunity to improve.
  - Perception that measurement is difficult.
  - If measurement activities are not integrated into work systems, they feel burdensome and like a distraction from the demands of daily business. Furthermore, measurement efforts that are not consolidated, or at least coordinated, across the organization often add unnecessary layers of complexity.
  - Measurement system design efforts are neglected.
  - In our experience, measurement is often addressed as an afterthought rather than carefully incorporated into organizational planning. Any initiative undertaken without a thoughtful planning process ultimately faces implementation challenges: measurement is no different.
  - Staff has little visibility for how measures are used.
  - Staff may not be supportive of measurement because they do not feel a connection to it or see how it can benefit them.

5.8 Integrity audits

Performance measurements should be scrutinized, just like other functions and processes. Financial indicators and the financial control and accounting system they are typically part of receive an annual audit by an external (third-party) firm. Non-financial strategic performance indicators do not consistently receive the same scrutiny. So how do managers know that these non-financial indicators are providing them with valid, accurate, and reliable information? Valid information here refers to face or content validity: does the indicator measure what it purports to measure? Reliable information means consistency in producing the same measurement output (i.e., indicator value) when identical performance conditions are repeatedly measured. Accuracy refers to how close the measurement output values are to the true performance values. By assuming that the indicators are providing valid, accurate, and reliable information, what assurance do managers have that their measurement systems are clearly understood, useful, and add value to the organization? A certain amount of financial measurement is a necessary part of doing business, for quarterly and annual SEC filings, reports to shareholders, or mandated by legislation in order to continue receiving government funding. The non-financial components of the measurement system are not typically mandated by legislation with the exception of compliance statistics like those reported to worker safety and environmental protection agencies. Organizations compelled to participate in supplier certification programs or achieve quality or environmental management systems certification may feel coerced to develop a rudimentary non-financial measurement system. However, they should realize that the return from developing a strategic performance measurement system is not compliance, but is the provision of useful information that adds value to the organization through better decision-making and support for implementation. After investing the time and resources to develop a strategic performance measurement system, organizations should periodically audit that system for validity, reliability, and accuracy and assess the system for continued relevance and value added.

It is beyond the scope of this chapter to describe the audit and assessment process in detail. The interested reader should refer to Coleman and Clark (2001). Figure 5.2 provides an overview of where the techniques suggested by Coleman and Clark can be applied to audit and assess the measurement process. "Approach" in the figure includes deciding on the extent of the audit and assessment, balancing previous efforts with current needs, and choosing among the variety of techniques available. The techniques in the figure are shown at the phases of the measurement and evaluation process where they are most applicable. Table 5.2 provides brief descriptions of these techniques and sources for additional information.

Organizations concerned with the resource requirements to develop, operate, and maintain a measurement system may balk at the additional tasking of conducting a comprehensive audit and assessment. Such organizations should, at a minimum, subject their measurement system to a critical review, perhaps using a technique as simple as "start, stop, or continue." During or immediately following a periodic review of performance (where the current levels of performance on each key indicator are reviewed and

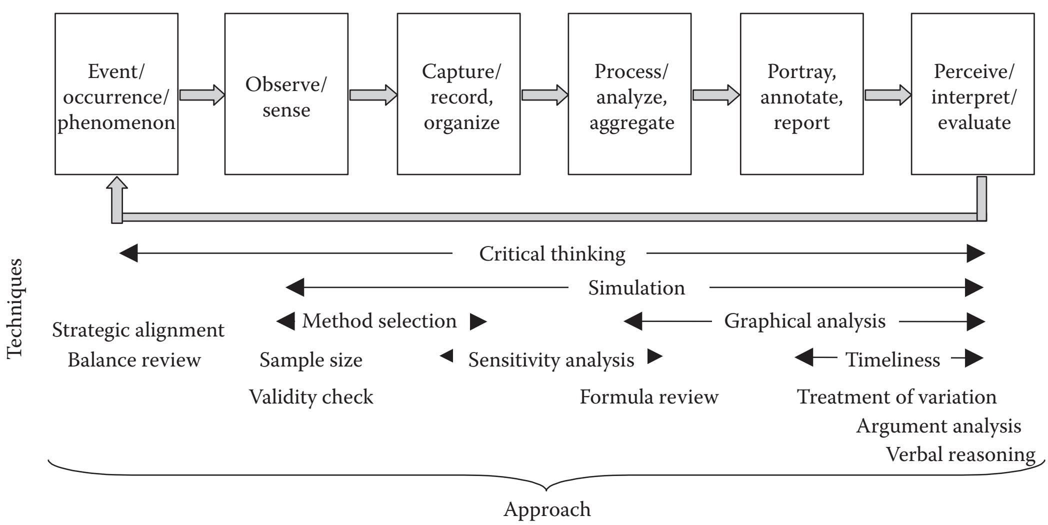

Figure 5.2 Auditing and assessing the measurement and evaluation process. (Adapted from Coleman, G.D. and Clark, L.A., A framework for auditing and assessing non-financial performance measurement systems, in *Proceedings of the Industrial Engineering Research Conference*, Dallas, CD-ROM, 2001.)

Table 5.2 Techniques Available for Auditing and Assessing Strategic Performance Measurement Systems

1. Strategic alignment—audit against the organization's priorities, implicit and explicit.
2. Balance review—assessment against the elements of one or more "balance" frameworks (e.g., Kaplan and Norton's Balanced Scorecard, Barrett's Balanced Needs Scorecard, Sink's Seven Criteria).
3. Critical thinking—scrutinizing for "faulty assumptions, questionable logic, weaknesses in methodology, inappropriate statistical analysis, and unwarranted conclusions" (Leedy, 2001, p. 36). Includes assessing the logic of the hierarchy of measures and the aggregation schemes. Assess value and usefulness by using Brown's (1996) or Sink's guidelines for the number of indicators used at one level in the organization.
4. Sample design—assessing sample design and the appropriateness of the generalizations made from these samples (i.e., external validity). This is more than an issue of sample size. "The procedure of stratification, the choice of sampling unit, the formulas prescribed for the estimations, are more important than size in the determination of precision" (Deming 1960, p. 28).
5. Validity check—auditing for evidence of validity. What types of validity have been established for these measures: face, content, construct, or criterion validity?
6. Method selection—assessment of the appropriateness of the method(s) chosen for the data being used. Includes choice of quantitative and qualitative methods. Might include assessment of the reliability of the methods. Internal validity might be addressed here.
7. Simulation—observing or entering data of known properties (often repeatedly), then comparing the output (distribution) of the measurement process against expectations.
8. Sensitivity analysis—varying input variables over predetermined ranges (typically plus and minus a fixed percent from a mean or median value) and evaluating the response (output) in terms of percentage change from the mean or median output value.
9. Formula review—comparison of the mathematical formulae to the operational and conceptual definitions of the measure. Also includes auditing of replications of the formulae to ensure consistent application.
10. Graphical analysis—at its simplest, plotting results and intermediate outputs to identify underlying patterns. In more advanced forms, may include application of statistical techniques such as individual and moving range charts (Wheeler, 1993). Assess any underlying patterns for possible impact on validity.
11. Timeliness—an assessment of the value of the information provided on the basis of how quickly the measured results reach someone who can directly use the results to control and improve performance. One simple technique is to track the lag time between occurrence and reporting of performance, then apply a user-based judgment of the acceptability of this lag.
12. Treatment of variation—graphical analysis is one technique for addressing variation. More importantly, how do the users of the measurement information perceive or react to variation in results? Assess available evidence of statistical thinking and the likelihood of interpreting noise as a signal or failing to detect a signal when present.
13. Argument analysis—"discriminating between reasons that do and do not support a particular conclusion" (Leedy and Ormrod, 2001, p. 36). Can be used to assess clarity with the Sink et al. (1995) technique described in Coleman and Clark (2001).
14. Verbal reasoning—"understanding and evaluating the persuasive techniques found in oral and written language" (Leedy and Ormrod, 2001, p. 36). Includes assessing the biases found in portrayal of performance information.

evaluated), the manager or management team using the measurement system should ask the following questions:

What should we start measuring that we are not measuring now? What information needs are currently unmet?
Which indicators that we are currently measuring should we stop measuring? Which are no longer providing value, are no longer relevant, or never met our expectations for providing useful information?
Which indicators should we continue to measure, track, and evaluate? If we were designing our measurement system from scratch, which of our current indicators would appear again?

Another less resource-intensive approach is to address the auditing and assessing of the measurement system as part of a periodic organizational assessment.

5.9 Organizational assessments: strategic snapshots of performance

Organizational assessments are a periodic snapshot form of strategic performance measurement. They are periodic in that they do not measure performance frequently: once a year to once every 5 or 10 years is common. They are snapshots because they reflect the organization's performance at a particular time and may not be fully evaluated until several weeks or months later. They are relatively comprehensive in scope, often measuring and evaluating all or most of the enterprise's activities and results, including the organization's measurement and evaluation system. Preparing for an organizational assessment may require a review of the organization's measurement system, and the assessment process will provide both direct and indirect feedback on the usefulness and value of the measurement system. Organizational assessments are used for conformity, to ensure the organization meets some standard (e.g., accreditation, certification), or for improvement and recognition where the organization is compared with a standard and provided feedback for improvement. Those exhibiting the highest levels of performance against the standard are recognized with an organizational award (e.g., Baldrige Award, State or Corporate Awards for Excellence, EFQM Excellence Award).

Organizational assessment typically begins with a self-study comparing the organization and its goals against an established standard (i.e., criteria or guidelines). The completed self-study is then submitted to a third party (i.e., the accreditation, registration, or award body) for review and evaluation. This third-party review begins with an evaluation of the self-study and is often, but not always, followed by a visit to the organization. The purpose of the visit is to validate and clarify what was reported in the self-study. The third party then renders a judgment and provides feedback to the organization. Depending on the specific application, the third-party judgment may result in substantial consequences for the organization (e.g., winning an award, receiving accreditation, or failure to do so). Ideally, the feedback from the third party is fed into the organization's improvement cycle, implemented, measured, and reflected in future plans and results.

Organizations that operate an ongoing improvement cycle and feed the results of the assessment into that cycle are likely to receive the greatest return on the investment from the resources required to complete the self-study and assessment. Particularly in situations where the organizational assessments occur several years apart, having an ongoing improvement process maintains the momentum and focus on what is important and should make preparing for future assessments easier. The improvement process translates assessment findings into plans, actions, and targets; applies resources; and then follows up with regular review of results and then new or updated plans, actions, and targets. While the overall improvement process should be management led, industrial engineers are often tasked as analysts and project managers to convert assessment findings into plans, actions, and results.

Organizations wishing to gain much of the benefit of a comprehensive assessment but concerned about the resource requirements should simply complete a five-page organizational profile, the preface of a Baldrige Award application (self-study) (Baldrige Performance Excellence Program, 2013, pp. 4–6). The organizational profile asks the organization to document its organizational environment, including product offerings, vision and mission, workforce, facilities, technologies, equipment, and regulatory requirements; its organizational relationships, including organization structure, customers and stakeholders, suppliers and partners; its competitive environment, including competitive position(s), competitiveness changes, and comparative data for evaluating performance; its strategic context in terms of key business, operational, social responsibility, and human resource challenges and advantages; and a description of its performance improvement system. For many organizations, particularly, smaller organizations and departments or functions within larger organizations, developing and collectively reviewing the organizational profile may provide more than 50% of the value of a complete organizational assessment. Too few management teams have developed consensus answers to the questions posed by the organizational profile. Developing the organizational profile as a team and keeping it current provides a key tool for providing organizational direction and furnishes an important input into the development and maintenance of the performance measurement system. Even organizations not interested in the Baldrige or other business excellence awards can use the profile as a resource for the development of management systems or the preparation of a self-study.

Organizational assessments, like other forms of performance measurement, should be subject to periodic audit and assessment. The reliability and validity of the results of organizational assessments are not as well investigated as we might like. Few, if any, of the organizations that offer or manage these assessments provide statistics showing they periodically evaluate the efficacy of their assessment processes. Researchers (Coleman et al., 2001, 2002; Coleman and Koelling, 1998; Keinath and Gorski, 1999; Van der Wiele et al., 1995) have estimated some of the properties associated with the scores and feedback received from organizational assessments. Their findings suggest that training the assessors (a.k.a. evaluators, examiners) reduces scoring leniency; however, their findings are less conclusive regarding the effect of training on interrater reliability and accuracy. Those interested in interpreting the variability observed among results from organizational assessments should consult the above-cited sources.

References

Akao, Y., Ed., *QFD: Quality Function Deployment*, Productivity Press, Cambridge, MA, 1990.

Baldrige Performance Excellence Program, *2011–2012 Baldrige Criteria for Performance Excellence*, National Institute of Standards and Technology, Gaithersburg, MD, 2011.

Baldrige Performance Excellence Program, *2013–2014 Criteria for Performance Excellence*, National Institute of Standards and Technology, Gaithersburg, MD, 2013.

Barrett, R., Liberating the corporate soul, in *Presentation to The Performance Center*, Arlington, VA, 1999a.

Barrett, R., *Liberating the Corporate Soul*, Fulfilling Books, Alexandria, VA, 1999b.

Becker, B., Huselid, M.A., and Ulrich, D., *The HR Scorecard: Linking People, Strategy and Performance*, Harvard Business School Press, Boston, 2001.
Bititci, U., Garengo, P., Dörfler, V., and Nudurupati, S., Performance measurement: Challenges for tomorrow, *Int. J. Manage. Rev.*, 14, 305–327, 2012.
Brassard, M. and Diane Ritter, GOAL/QPC, *The Memory Jogger*, Methuen, MA, 1985.
Brown, M.G., *Keeping Score: Using the Right Metrics to Drive World-Class Performance*, AMACOM Books, New York, 1996.
Brown, M.G., *Winning Score: How to Design and Implement Organizational Scorecards*, Productivity Press, Portland, OR, 2000.
Busi, M. and Bititci, U.S., Collaborative performance management: Present gaps and future research, *Int. J. Prod. Perform. Manage.*, 55, 7–25, 2006.
Coleman, G.D. and Clark, L.A., A framework for auditing and assessing non-financial performance measurement systems, in *Proceedings of the Industrial Engineering Research Conference*, Dallas, 2001 (CD-ROM).
Coleman, G.D., Costa, J., and Stetar, W., The measures of performance: Managing human capital is a natural fit for industrial engineers, *Ind. Eng.*, 36, 40–44, 2004.
Coleman, G.D. and Koelling, C.P., Estimating the consistency of third-party evaluator scoring of organizational self-assessments, *Qual. Manage. J.*, 5, 31–53, 1998.
Coleman, G.D., Koelling, C.P., and Scott Geller, E., Training and scoring accuracy of organisational self-assessments, *Int. J. Qual. Reliab. Manage.*, 18, 512–527, 2001.
Coleman, G.D., Van Aken, E.M, and Shen, J., Estimating interrater reliability of examiner scoring for a state quality award, *Qual. Manage. J.*, 9, 39–58, 2002.
Cooper, W.W., Seiford, L.M., and Zhu, J., Eds., *Handbook on Data Envelopment Analysis*, Springer, New York, 2004.
Deal, T.E. and Kennedy, A., *Corporate Cultures*, Addison-Wesley, Reading, MA, 1982.
Delbecq, A.L., Van De Ven, A.H., and Gustafson, D.H., *Group Techniques for Program Planning: A Guide to Nominal Group and Delphi Processes*, Green Briar, Middleton, WI, 1975.
Deming, W.E., *Sample Design in Business Research*, Wiley, New York, 1960.
Deming, W.E., *The New Economics for Industry, Government, and Education*, MIT Center for Advanced Engineering Study, Cambridge, MA, 1993.
Drucker, P.F., *The Practice of Management*, Harper & Row, New York, 1954.
Drucker, P.F., The next information revolution, *Forbes ASAP*, 47–58, 1998.
Farris, J.A., Van Aken, E.M., Letens, G., Chearksul, P., and Coleman, G., Improving the performance review process: A structured approach and case application, *Int. J. Oper. Prod. Manage.*, 31, 376–404, 2011.
Garvin, D.A., What does "product quality" really mean?, *Sloan Manage. Rev.*, 26, 25–43, 1984.
Hamel, G., *Leading the Revolution*, Harvard Business School Press, Cambridge, MA, 2002 (Revised).
Hamel, G. and Prahalad, C.K., *Competing for the Future*, Harvard Business School Press, Cambridge, MA, 1996.
Ishikawa, K., *What Is Total Quality Control? The Japanese Way* (trans. Lu, D.J.), Prentice-Hall, Englewood Cliffs, NJ, 1985.
Kaplan, R.S. and Norton, D.P., The balanced scorecard: Measures that drive performance, *Harvard Bus. Rev.*, 70, 71–79, 1992.
Kaplan, R.S. and Norton, D.P., Using the balanced scorecard as a strategic management system, *Harvard Bus. Rev.*, 74, 75–85, 1996.
Keinath, B.J. and Gorski, B.A., An empirical study of the Minnesota Quality Award Evaluation Process, *Qual. Manage. J.*, 6, 29–38, 1999.
Kilmann, R., *Managing Beyond the Quick Fix*, Jossey-Bass, San Francisco, 1989.
Kilmann, R. and Saxton, M.J., *Kilmann-Saxton Culture Gap Survey*, XICOM, Tuxedo, NY, 1983.
Kizilos, T., Kratylus automates his urnworks, *Harvard Bus. Rev.*, 62, 136–144, 1984.
Kubiak, T.M. and Benbow, D.W., *The Certified Six Sigma Black Belt Handbook*, 2nd ed., Quality Press, Milwaukee, WI, 2009.
Leedy, P.D. and Ormrod, J.E., *Practical Research: Planning and Design*, Merrill Prentice Hall, Upper Saddle River, NJ, 2001.

Mallak, L.A., Bringelson, L.S., and Lyth, D.M., A cultural study of ISO 9000 certification, *Int. J. Qual. Reliab. Manage.*, 14, 328–348, 1997.

Mallak, L.A. and Kurstedt, H.A., Using culture gap analysis to manage organizational change, *Eng. Manage. J.*, 8, 35–41, 1996.

Medina-Borja, A.M., Pasupathy, K.S., and Triantis, K., Large-scale data envelopment analysis (DEA) implementation: A strategic performance management approach, *J. Oper. Res. Soc.*, 58, 1084–1098, 2006.

Miller, D., Profitability = productivity + price recovery, *Harvard Bus. Rev.*, 62, 145–153, 1984.

Muckler, F.A. and Seven, S.A., Selecting performance measures: "Objective" versus "subjective" measurement, *Hum. Factors*, 34, 441–455, 1992.

Neely, A., The performance measurement revolution: Why now and what next?, *Int. J. Oper. Prod. Manage.*, 19, 205–228, 1999.

Nørreklit, H., The balanced scorecard: What is the score? A rhetorical analysis of the balanced scorecard, *Account. Org. Soc.*, 28, 591–619, 2003.

Peters, T.J. and Waterman, R.H., Jr., *In Search of Excellence: Lessons from America's Best-Run Companies*, Warner Books, New York, 1982.

Pfeffer, J. and Sutton, R.I., *The Knowing-Doing Gap: How Smart Companies Turn Knowledge into Action*, Harvard Business School Press, Boston, 2000.

Pineda, A.J., Productivity measurement and analysis (module 4), in *Productivity and Quality Management: A Modular Programme*, Prokopenko, J. and North, K., Eds., International Labor Organization, Geneva, 1996.

Rucci, A.J., Kirn, S.P., and Quinn, R.T., The employee-customer-profit chain at Sears, *Harvard Bus. Rev.*, 76, 82–97, 1998.

Sink, D.S., *Productivity Management: Planning, Measurement and Evaluation, Control and Improvement*, Wiley, New York, 1985.

Sink, D.S. and Tuttle, T.C., *Planning and Measurement in Your Organization of the Future*, IIE Press, Norcross, GA, 1989.

Sumanth, D.J., *Total Productivity Management: A Systematic and Quantitative Approach to Compete in Quality, Price, and Time*, St. Lucie Press, Boca Raton, FL, 1998.

Thompson, A.A., Jr. and Strickland, A.J., III, *Strategic Management: Concepts and Cases*, 13th ed., Irwin/McGraw-Hill, Boston, 2003.

Thor, C.G., *Designing Feedback*, Crisp Publications, Menlo Park, CA, 1998.

Tufte, E.R., *Visual and Statistical Thinking: Displays of Evidence for Decision Making*, Graphics Press, Cheshire, CT, 1997a.

Tufte, E.R., *Visual Explanations: Images and Quantities, Evidence and Narrative*, Graphics Press, Cheshire, CT, 1997b.

Van Der Wiele, T., Williams, R., Kolb, F., and Dale, B., Assessor training for the European Quality Award: An examination, *Qual. World Tech. Suppl.*, 1995, 53–62, 1995.

Wheeler, D.J., *Understanding Variation: The Key to Managing Chaos*, SPC Press, Knoxville, TN, 1993.

chapter six

Industrial engineering applications in the construction industry

Lincoln H. Forbes*

Contents

* Also acts as an Adjunct Professor in the College of Engineering, Florida International University, Miami, FL 33199, USA.

6.1 *Introduction*

The purpose of this chapter is to provide an overview of the possible applications of industrial engineering (IE) techniques in construction. Due to space restrictions and the large number of techniques available, a very limited selection of examples is presented. Readers are encouraged to do further reading of the sources provided in the references.

The Construction Industry has traditionally been one of the largest industries in the United States. As reported by the Bureau of Labor Statistics (BLS), U.S. Department of Labor, the value of construction put in place in 2003 was $916 billion, representing 8.0% of the gross domestic product. The industry employed approximately 6.9 million people in 2003. By its very nature, construction activity in the United States has not been subjected to the trend toward outsourcing that has plagued both the manufacturing and service industries. The BLS report titled "*State of Construction 2002–2012*" forecasts that 58.4% of U.S. jobs will be construction-related at the end of that decade. Yet, although other industries have blazed a trail to higher levels of quality and performance, the majority of construction work is based on antiquated techniques.

The potential for savings and productivity improvement is immense. Studies have pointed to typical losses in construction projects in the range of 30%; were this projected to the nation's annual total, over 200 billion may be wasted in a variety of ways. Mistakes, rework, poor communication, and poor workmanship are part of an ongoing litany of deficiencies that seem to be accepted as being a natural part of construction activity. Safety is a major national concern. Construction has an abysmally poor safety record, worse than virtually all other industries.

6.1.1 Categories of construction

In order to understand how IE techniques can be applied to the construction industry, it is helpful to understand that environment; it is truly diverse, so much so that its participants have found it easy to rely on such clichés as "the industry is like no other," "no two projects are alike," to maintain the status quo in which long-established management traditions are seen as an arcane art that others cannot understand fully.

The BLS refers to three major headings: General Building Contractors SIC Code 15, Heavy Construction (except building) SIC Code 16, and Special Trade Contractors SIC 17. These are further subdivided into 11 SIC Code headings that include:

- Commercial building construction: offices, shopping malls
- Institutional construction: hospitals, schools, universities, prisons, etc.
- Residential: housing construction, including manufactured housing
- Industrial: warehouses, factories, and process plants
- Infrastructure: road and highway construction, bridges, dams, etc.

Who are the parties involved in construction?

- Owners, who originate the need for projects and determine the locations and purpose of facilities.
- Designers—they are usually architects or engineers (electrical, mechanical, civil/structural), who interpret the owner's wishes into drawings and specifications that may be used to guide facility construction. In the design-build (DB) process, they may be part of the construction team.
- Constructors—they are contractors and subcontractors who provide the workforce, materials, equipment and tools, and provide leadership and management to implement the drawings and specifications to furnish a completed facility.
- Construction trades, represented by unions.
- Consumer advocates and building owners.
- The legal industry.
- Developers.
- Major suppliers.
- Code enforcement professionals.
- Financial institutions—banks, construction financial organizations.
- Safety professionals.

6.1.2 Construction delivery methods

Several methods are available for carrying out construction projects. Design-bid-build (DBB) is the most traditional method of project delivery. Typically, a project owner engages a design organization to conduct planning, programing, and preliminary and detailed design of facilities. The final design and specifications are used to solicit bids from contractors. A contractor is hired with a binding contract based on the owner's drawings and specifications. Because of the linear nature of this process, several years may elapse between project conceptualization and final completion.

Design-build involves a contractor and designer working as a combined organization to provide both design and construction services. The owner engages a design professional to do a limited amount of preliminary project planning, schematic design, cost, and schedule proposals. DB firms subsequently compete for a contract based on the owner's preliminary information. The selected DB may commence construction while completing the final design. This concurrent engineering approach significantly reduces the duration of each project.

Engineer–procure–construct (EPC) contracts are similar to design build; this type of delivery involves a single organization providing engineering, procurement, and construction. It is most appropriate for engineering-based projects such as construction of manufacturing facilities or large municipal projects.

Construction management (CM) involves coordination and management by a CM firm of design, construction activities. The owner may elect to pay a fee for these services. CM at risk, on the other hand, involves the assumption of risk by the contractor for carrying out the construction through its own forces. Other types of delivery systems may be based on a combination of the foregoing systems. Overall, the methods have advantages and disadvantages that are best identified through systematic analysis.

6.2 Industrial engineering applications

There are several areas in the construction industry where IE techniques may be applied. The techniques are as follows:

- Ergonomics/human factors
- Value engineering
- Work measurement
- The learning curve
- Quality management (QM)
- Productivity management
- Continuous improvement
- ISO 9000
- Cycle time analysis
- Lean methods
- Supply-chain management (SCM)
- Automation/robotics
- Radio frequency identification (RFID)
- Safety management
- Systems integration
- Simulation
- Quality function deployment
- Facilities layout
- Operations research and statistical applications
- Sustainable construction

6.2.1 Ergonomics/human factors

The study and redesign of construction workspace using traditional and modern IE tools could increase efficiency and minimize on-the-job injuries and worker health impacts. Unlike factories, construction workspace constantly changes in geometry, size, location and type of material, location of work, location of material handling equipment and other tools, etc. These create new and challenging research opportunities. In addition, significant environmental impacts result from construction-related activities. Safety engineering approaches and industrial ecology tools such as life-cycle analysis may be developed to define and measure the impacts of different designs for workspaces and constructions.

Construction workers use a wide assortment of tools and equipment to perform construction tasks. Especially in cases where such aids are used for prolonged periods of time, workers' effectiveness and capacity to work with high levels of concentration, ergonomics are a major concern. Workers cannot be expected to "build in" quality in constructed facilities if they are subjected to awkward positions and excessive physical stress caused by tools and equipment that are difficult to use.

The significance of ergonomics in the construction environment is evident from a study conducted by the Associated General Contractors (AGC) of California to examine ergonomics-related costs. Their findings are:

- Related Workers Comp Insurance claims had increased by up to 40% for many construction companies.
- Financial returns due to ergonomic business strategy—80% of the companies that had incorporated ergonomics-based methods reported improvements.
- Of 24 companies that measured for productivity, 100% reported improvements in cases where ergonomics-related concerns were addressed.

6.2.1.1 Tool and equipment design

Much research has yet to be done in the design of construction-oriented tools and equipment. The factors that may cause fatigue include weight, size, vibration, and operating temperature. Work-related musculoskeletal disorders (WRMSDs) generally include strains, sprains, soft tissue, and nerve injuries; they are cumulative trauma disorders and repetitive motion injuries. The construction workers who are at highest risk for these disorders are carpenters, plumbers, drywall installers, roofers, electricians, structural metal workers, carpet layers, tile setters, plasterers, and machine operators.

The top five contributory risk factors are as follows:

Working in a specific posture for prolonged periods, bending or rotating the trunk awkwardly, working in cramped or awkward positions, working after sustaining an injury, and handling heavy materials or equipment.

The use of a shovel is a very typical example of the labor-intensive material handling activities that are routinely carried out on construction projects. This activity requires workers to bend over, apply force to a shovel in different planes, and rotate the trunk in a flexed position. Such movements impose biomechanical stress which may impose cumulative trauma risk. Freivalds (1986) studied the work physiology of shoveling tasks and identified the shovel design parameters that would increase task efficiency. Friedvald's two-phase experimental study addressed the following parameters:

- The size and shape of the shovel blade
- The lift angle
- Shovel contours—hollow and closed-back design
- Handle length
- Energy expenditure
- Perceived exertion
- Low-back compressive forces

The recommended shovel design is as follows:

- A lift angle of approximately 32°
- A hollow-back construction to reduce weight
- A long tapered handle
- A solid socket for strength in heavy-duty uses
- A large, square point blade for shoveling
- A round, point blade for digging, with a built-in step for digging in hard soil

6.2.1.2 Ergonomics applications in structural ironwork

The BLS reports that construction trade workers experience higher rates of musculoskeletal injuries and disorders than workers in other industries: 7.9 cases per 100 equivalent workers as compared with the industry average of 5.7 per 100 (Bureau of Labor Statistics, 2001). In overall injuries, construction workers registered 7.8 vs. the industry average of 5.4. Observations by Holstrom et al. (1993), Guo et al. (1995), Kisner and Fosbroke (1994), and others point to a lack of studies in ergonomics, presumably because of high task variability, irregular work periods, changing work environments, and the transient nature of construction trades. As pointed out by Forde and Buchholz (2004), each construction trade and task represents a unique situation; the identification and application of prevention measures, tools and work conditions is best derived from trade and task-specific studies. This approach is the most likely to minimize the incidence of construction trades' WRMSDs.

By way of illustration, Forde and Buchholz (2004) studied construction ironworkers to identify mitigating measures in that group. Construction ironwork refers to outdoor work (not shop fabrication) as four specialties—the erection of structural steel (structural ironwork [SIW]), placement of reinforcing bars (rebars) (reinforcing ironwork [RIW]), ornamental ironwork (OIW), and machinery moving and rigging (MMRIW).

Previous studies determined that construction ironwork involves lifting, carrying and manipulating of heavy loads, maintaining awkward postures in cramped quarters, working with arms overhead for extended periods, using heavy, vibrating pneumatic tools, and extensive outdoor exposure in temperature and weather extremes.

Forde and Buchholz (2004) made the following observations and recommendations on the various categories of ironwork:

- *Machinery moving/rigging*. The erection of equipment such as a crane involves the pushing and pulling of large and heavy segments, and lining them up for bolting together. During an 8-h shift, this activity was observed to require 1.3 h of significant whole-body exertion. Workers in this scenario are most susceptible to overexertion of the back, legs, and shoulders.

- *Ornamental ironwork*. This work was observed to require arms to be above the shoulder level 21% of the time. Trunk flexion or twisting and side bending were observed 23% of the time.

These percentages indicate a high risk of overexertion of the involved muscle groups. Industrial engineers should review the work methods to increase the amount of preassembly at workbench height.

- *Reinforcing ironwork*. The preparation of reinforcement cages and tying of rebars were seen to cause nonneutral trunk postures up to 50% of the time. The handling of heavy loads (50 lb or greater) was observed to occur for 1.9 h of an 8-h shift, representing significant long-term risk. A 2004 study by Forde and Buchholz identified a need to improve the design of hand tools used for securing rebars. Such redesign would reduce nonneutral hand/wrist postures such as flexion, extension, and radial and ulnar deviation. These postures put construction workers at risk of repetitive motion injuries.

6.2.1.3 Auxiliary handling devices

A number of research studies have shown that construction workers have suffered back, leg, and shoulder injuries because of overexertion resulting from stooped postures, performing manual tasks above shoulder level, and the lifting of heavy objects. Such overexertion and injuries reduce worker productivity and may negatively affect the timeliness and profitability of construction projects. The use of auxiliary handling devices may reduce the degree of overexertion experienced by construction workers, and enhance productivity. Sillanpaa et al. (1999) studied the following five auxiliary devices:

- Carpet wheels
- A lifting strap for drain pipes
- A portable cutting bench for molding
- A portable storage rack
- A portable cutting bench for rebars.

The survey subjects utilized these devices to carry out typical construction tasks, such as carrying rolls of carpet, mounting drain pipes, cut pieces of molding, and fashioned rebars. The results of the study were mostly positive but mixed, pointing to the need for further research. The auxiliary devices were found to reduce the muscular load of some subjects, but others experienced an increased load because of differences in anthropometric dimensions, work modes, and level of work experience.

6.2.1.4 Drywall hanging methods

Drywall lifting and hanging are extensively conducted in both residential and commercial building construction; drywall board has become the standard for interior wall panels. It is the standard for surfacing residential ceilings. Workers are required to handle heavy and bulky drywall sheets and assume and maintain awkward postures in the course of performing installation work. These activities often cause muscle fatigue and lead to a loss of balance; studies have identified drywall lifting and hanging tasks as causing more fall-related injuries than any other tasks. Pan et al. (2000) studied 60 construction workers to identify the methods resulting in the least postural stability during drywall lifting and hanging tasks.

The subjects' instability was measured using a piezoelectric-type force platform. Subjects' propensity for loss of balance was described by two postural-sway variables (sway length and sway area) and three instability indices (PSB, SAR, and WRTI). The study was a randomized repeated design with lifting and hanging methods for lifting and hanging randomly assigned to the subjects. ANOVA indicated that the respective lifting and hanging methods had significant effects on two postural-sway variables and the three postural instability indices.

The recommended methods were:

- Lifting drywall sheets horizontally with both hands positioned on the top of the drywall causes the least postural sway and instability.
- Hanging drywall horizontally on ceilings produces less postural sway and instability than vertically.

6.2.2 *Value engineering*

Value engineering (VE) is a proven technique for identifying alternative approaches to satisfying project requirements, while simultaneously lowering costs. It is a process of relating the functions, the quality, and the costs of a project in determining optimal solutions (Dell'Isola, 1988). In the construction environment it involves an organized multidisciplined team effort to analyze the functions of systems, equipment, facilities, processes, products, and services to achieve essential functions at the lowest possible cost consistent with the customer's requirements while improving performance and quality requirements. The multidisciplined approach of the IE is well suited to driving and facilitating the VE process. The IE can be especially valuable in facilitating a multidisciplinary group of design and construction professionals in brainstorming, generating ideas, and in conducting life-cycle analysis for the comparison of alternatives.

Some client organizations, such as government agencies, share the savings derived from VE with the contractor; the ratio varies with the respective type of contract.

Private contractors are generally highly motivated to develop improvements to tasks or projects, because of the financial benefit of lowering their costs; lower costs translate to higher profits.

There are several examples of savings in construction value engineering:

The U.S. Army Corps of Engineers has been using VE principles since 1964; in 2001, the Corps saved $90.78 million in its Civil works programs and has also realized at least $20 for each $1 spent on VE. $421 million in life-cycle cost was saved on a criminal court complex in New York City using VE. At the Bayou Bonfouca project in Louisiana, capital savings of $200,000 were obtained, and operations and maintenance costs were even greater at $4.4 million over a 2-year period.

Value engineering was successfully applied in a project at the Port of San Diego General Services Facilities building. The new structure comprised 45,200 sq. ft of administrative offices and maintenance shops at a cost of $8.9 million. A VE consultant was hired for the project. The building cost was reduced by 10%. In addition, the VE application placed a high priority on energy efficiency. The design was modified to emphasize the use of natural convection ventilation in shop areas vs. forced air, and specialized lighting/controls were selected to reduce energy consumption. Consequently, energy costs were reduced by 10%.

Kubal (1994) points out that while VE is beneficial during the design stages of a project, it can be most effective during the preconstruction phase because it facilitates both product and process improvements. Therefore, it should be perceived not just as a cost reduction exercise, but as a means of improving the entire construction process.

Design for manufacture and assembly (DFMA) techniques may be used to supplement VE activities; DFMA involves the review of designs to identify the optimal choice of materials, component design, fabrication, and assembly for the most cost-effective and functional solution. DFMA is carried out with the participation of a multidisciplinary team—whereas in manufacturing environments the team includes manufacturing engineers, shop floor mechanics, suppliers' representatives and specialists in maintainability and reliability studies, construction projects would include building design engineers, architects, contractors, and maintenance personnel.

A typical VE project may involve seven phases:

Team selection. A VE team leader supervises a number of team members. These individuals should preferably be construction professionals who are generalists and specialists; the team leader should seek out flexible individuals who are willing to participate in a group activity. The team members should be trained in the VE process.

Information gathering. Team members gather information on both technical and cost issues relating to a project, using available documents; the team VE leader assembles the information and shares it with the entire team.

Brainstorming. This phase involves creative thinking to identify alternatives for carrying out a project. Experienced team members may recommend innovative approaches for conducting a project. The brainstorming phase is expected to generate many ideas without judgment. The original design is the point of reference for the alternatives that are generated.

Evaluation of alternatives. Each alternative is reviewed carefully to determine its feasibility. Cost benefit or life-cycle cost analysis may be conducted in order to rank the possible solutions in order of importance. This ranking may be based on cost, and also on, ease of implementation.

Recommendation of alternatives. The team leader reports on all the alternatives to the team, then selects the most appropriate ones for the client/owner. The savings derived may be in the range of 5%–30% of initial project cost estimates.

Implementation. The contractor implements the selected alternatives and the savings are divided between the owner and the contractor. The method of division is generally dictated by the form of contract. In the case of U.S. Government contracts, for example, Federal Acquisition Regulations (FAR) advocate the use of VE to reduce project costs. It also prescribes the types of savings and sharing for each type of project; the ratios used for various contract types.

As described by Adrian and Adrian (1995), the VE process matches the worth and cost of building elements; aesthetically pleasing features should not represent a significantly higher percentage of building cost than those attributes that the owner considers most valuable. For example, it is not uncommon for facilities to be built with brass hardware and marble floors, yet lack adequate service access to HVAC equipment. The VE technique is most effective when applied to the design phases of a project, when the influence on cost is greatest.

Optimization of projects with VE. Typical factors to consider in optimizing construction projects are:

- The intended purposes and functions for a project/facility.
- A clear understanding of the owner/client's needs.
- The perceived value to users and aesthetic appeal.
- Architectural systems and finishes and the specified conditions for their operation.
- Structural systems and materials—to maintain the integrity of a project/facility under all design conditions.
- Electrical, lighting, and communications systems—adequate and reliable operation is required.
- HVAC, plumbing, gas, and other systems to maintain a comfortable environment for users.
- Fire protection systems for detection and fire-fighting, adequate means of egress in case of emergencies.
- The constructability of a facility—the proposed construction methods and the projected time frames.
- The maintainability of a facility, the maintenance requirements, and the replacement cycle for components (HVAC, lighting devices, flooring materials, etc.).
- The expected return on investment for the owner/client.

In applying the VE process to building systems and components, the following steps may be used:

1. Identify functions
2. Estimate the value of each function
3. List the components
4. Determine component costs
5. Identify component functions
6. Calculate the cost per function
7. Evaluate and modify the proposed design

A VE team is staffed by knowledgeable individuals—designers, maintenance staff, etc., who understand the consequences of their decisions. They are also trained in the VE process, and participate in steps 1–7 given above.

6.2.3 Work measurement

Work measurement techniques can help to increase construction productivity. Whereas standard work times are often used by the industry, these standards need to be reviewed and updated. Industrial engineers can tailor these standards to specific projects to reflect the logistics of the work site and also adjust the standards to represent methods improvement. The more accurate the information that is available on work standards, the better construction managers can conduct the preplanning of projects and exert greater control over the costs and schedules of these projects. Many construction standards need to be reengineered to reflect the use of technology in work processes. Methods time measurement (MTM) can be used to develop engineered standards.

Methods time measurement is based on the concept that a method must first be developed, elemental steps defined, and standard times developed. The standard must be based on the average times necessary for trained experienced workers to perform tasks of prescribed quality levels, based on acceptable trade practices. This approach is most practical with repetitious tasks.

In the MTM system, operations are subdivided into tasks; tasks are further reduced to individual body movements such as reaching, grasping, applying pressure, positioning, turning, and disengaging. Other movements include eye travel and focus and body, leg, or foot motions. Each body movement is subdivided into individual actions, such as reaching 2 in., grasp, apply pressure, turn, etc. Each action is assigned a standard time stated in time measurement units (TMU).

Examples of Methods—Time Measurement Application Data (All Times Include a 15% Allowance)

Activity	TMU	
Reach 2^2	4	
Grasp (simple)	2	
Turn	6	
Regrasp	6	ITMU = 0.00001 H
Look (eye time)	10	= 0.006 min
Leg motion	10	= 0.036 sec
Kneel on one knee	35	
Arise	35	

In applying the MTM system (or any other standardized measurement system) it cannot be overemphasized that an appropriate method must first be established that can be successfully applied by the average, trained worker at definitive quality levels. The effect of the learning curve should also be considered when establishing work standards to ensure that repetition does not render the task times excessively long.

6.2.4 The learning curve

A learning curve is the phenomenon demonstrated by the progressive reduction in the time taken by an individual, or by a team to perform a task or a set of tasks on a repetitious basis. The individuals performing the task or project become more proficient with each repetition; the observed improvement serves as a motivator and a learning tool resulting in successively shorter performance times.

The learning curve is represented by an equation of the form

$$T_n = T_1(n)^{**}(-a)$$

where T_n is the time for the nth cycle, T_1 the time for the first cycle, n the number of cycles, and a a constant representing the learning rate. This equation produces a hyperbolic curve.

In order to determine the learning rate of a given activity, time study may be applied to a worker who is performing the task. For example, masons installing concrete blocks to form a wall would be timed as they perform successive iterations of the process.

The learning curve can be applied to construction projects. It can be highly relevant in repetitious projects such as housing construction, but the success of this application requires the IE to understand that interruptions to the construction process limit its use. Examples of such interruptions include prolonged shutdowns and Christmas holidays.

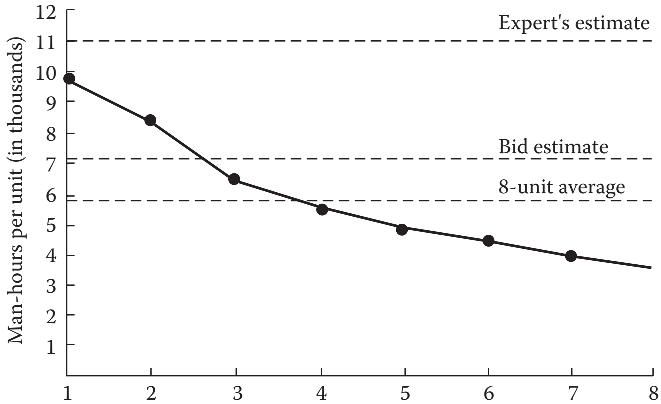

Figure 6.1 Progression of learning curve.

Also, construction tasks are often varied and nonrepetitive, so the IE has to apply the concept very judiciously. On-site managers who understand the learning curve rates for different types of tasks can improve work performance by selecting alternative work methods, especially with less experienced crafts persons.

Oglesby et al. (1989) identified three distinct phases: (1) when construction crews are familiarizing themselves with a process; (2) when a routine is learned so that coordination is improved; and (3) a deliberate and continuing effort to improve with successive iterations of the process.

Oglesby et al. (1989) estimated that learning curves for construction typically fall in the 70%–90% range.

The curve below represents a project involving the installation of a number of generator units. The expert's estimate for carrying out this work was 11,000 man-hours. The contractor's bid was lower, i.e., 7200 h per unit. It is unlikely that a bid based on the expert's estimate would have been successful. The use of the learning curve allowed the contractor to complete the project at an even lower level of man-hours, i.e., 5900 h per unit. By using the benefit of the learning curve, the contractor was able to reduce labor hours by 1200 × 8 = 9600 h over eight installations. This savings could translate directly to an increased profit margin. This profit margin is represented by the difference in area under the bid estimate line and the 8-unit average line (Figure 6.1).

6.2.4.1 Example—learning curve calculations

A construction crew is carrying out a repetitious task. The first cycle takes the crew 5 h to complete. The third cycle takes the crew 4 h. Learning rate can be calculated by

$$i = 1 \text{ (first cycle)}$$

$$j = 3 \text{ (third cycle)}$$

$$r = 5 \text{ h}$$

$$s = 4 \text{ h}$$

$$r/s = (j/i)^n$$

$$5/4 = (3/1)^n$$

$$1.25 = 3^n$$

$$n \log 3 = \log 1.25$$

$$n = \log 1.25/\log 3 = = 0.203$$

$$\text{Learning rate} = 2^n = 1/2^n = 1/1.151 = 0.868$$

$$\text{Learning rate} = 86.8\%$$

How long should it take to complete the fourth cycle of the task?

$$i = 1$$

$$j = 4$$

$$r = 5$$

$$s = ?$$

From above, learning rate = 86.8%, $n = 0.203$

$$r/s = (j/i)^n$$

$$5/s = (4)0.203 = 1.3205$$

$$s = 5/1.3205 = 3.79 \text{ h}$$

The fourth cycle takes 3.79 h.

6.2.5 *Quality management*

Total quality is an approach of doing business that attempts to maximize the competitiveness of an organization through the continual improvement of the quality of its products, services, people, processes, and environments (Goetsch and Davis, 2000).

Historically, the Japanese were among the first to apply quality improvement approaches in construction on a large scale, although they did not embrace this concept until the oil crisis of 1973. Prior to this, they thought that the construction industry was inappropriate for the application of total quality control (TQC), because of the inherent variability in projects and the difficulty in defining "acceptable quality." Takenaka Komuten Company, the sixth largest in Japan, had their formerly impeccable safety and quality image tarnished by the failure of a sheet piling system in Okinawa, in 1975, and embarked on a quality control (QC) program. They were followed by Shimizu Construction Company, the second largest in Japan, that established a QC program in 1976, and by Kajima Corporation, the third largest, in 1978. Subsequently, several U.S. companies have adopted TQC programs and the more familiar total quality management (TQM) programs used by U.S. manufacturers.

In 1992, the Construction Industry Institute (CII) published Guidelines for Implementing Total Quality Management in the Engineering and Construction Industry. Their research studies confirm that TQM has resulted in improved customer satisfaction, reduced cycle times, documented cost savings, and more satisfied and productive workforces (Burati and Oswald, 1993).

6.2.5.1 Benefits of TQM

The application of TQM principles can benefit design and construction organizations in many ways. These include

- Survival in an increasingly competitive world
- Improved levels of customer service
- Reduced project durations and costs
- Improvement of the overall quality and safety of facilities
- Better utilization of employees' skills/talents and increased quality orientation
- Increased profitability

6.2.5.2 Foundations of TQM

Total quality management is based on the total quality concept, which involves everyone in an organization in an integrated effort toward improved performance at each level (Goetsch and Davis, 2003).

It integrates fundamental management techniques, improvement efforts, in a disciplined approach toward continual process improvement. Total quality has the following characteristics: it is driven by an organizational strategy and unity of purpose, an internal and external customer focus, obsession with quality, scientifically based decision making and problem solving, continuous process improvement, long-term commitment, teamwork, employee involvement and empowerment, and education and training.

While total quality approaches have been highly beneficial to the manufacturing and service industries, they have had limited application in the construction environment. The construction industry has been heavily steeped in the traditional ways of executing projects and its constituents—designers and constructors, have been reluctant to make a necessary cultural and behavioral change to adopt total quality approaches.

Top management and senior management are generally preoccupied with short term, project by project profitability, and not with long-term quality-based strategies.

Although organizations have adopted a wide variety of quality improvement programs, these programs are based on the concepts advocated by the total quality pioneers. The most highly acknowledged pioneers are W. Edwards Deming, Joseph M. Juran, and Philip B. Crosby. Armand V. Feigenbaum and Japanese experts Kaoru Ishikawa and Shigeo Shingo were also major contributors to the quality improvement philosophy.

Deming has emerged as the influential and durable proponent of QM in the United States and is best known for the Deming cycle, his 14 points, and the seven deadly diseases.

The 14 points are summarized as:

1. Develop a program of constancy in purpose
2. Adopt this new program and philosophy
3. Stop depending on inspection to achieve quality—build in quality from the start
4. Stop awarding contracts on the basis of low bids
5. Improve continuously and forever the system of production and service
6. Institute training on the job
7. Institute leadership

8. Drive out fear so everyone may work efficiently
9. Eliminate barriers between departments so that people can work as a team
10. Eliminate slogans, targets, and targets for the workforce—they create adversarial relationships
11. Eliminate quotas and management by objectives
12. Remove barriers that rob people of pride of workmanship
13. Establish rigorous programs of education and self-improvement
14. Make the transformation everyone's job.

Juran is known for several quality contributions:

- Three basic steps to progress
- Ten steps to quality improvement
- The quality trilogy

Ishikawa is credited with the development/adaptation of seven quality tools:

- Pareto charts
- Cause and effect diagrams
- Scatter diagrams
- Check sheets
- The histogram
- Stratification
- Control charts

6.2.5.3 Obstacles to TQM

There are many obstacles to the application of TQM in the construction environment, and industrial engineers can help the industry to overcome these concerns:

1. Measuring results is difficult (Shriener et al., 1995), whereas Deming (1991) advocate that measurement is a critical element in quality improvement efforts. The concept of construction performance does not emphasize productivity and quality initiatives. The work of many researchers has revealed an industry tendency to measure performance in terms of the following: completion on time, completion within budget, and meeting construction codes. Very little attention has been directed to owner satisfaction as a performance measure.
2. *The industry has a crisis orientation.* Significant changes have been sparked primarily by catastrophes of one kind or another. Major revisions were made in U.S. engineering codes after the failure of a structure in the Kansas City Hyatt Regency Hotel. Hurricane Andrew devastated Dade County, Florida, in August 1992, resulting in a major scrutiny of building codes and their enforcement. It is probable that with sufficient attention to quality at the front end, more building failures might be avoidable.
3. *Poor communication.* Communication tends to be via the contract. Essentially, the designer is paid to produce a design expressed in the form of specifications and drawings. The contractor is expected to use these as a means of communication, and produce the completed facility. This communication often does not work as well as it

should. Cross-functional communication must include subcontractors and suppliers to solve quality problems.

4. There are large gaps between expectations and results as perceived by construction owners. Symbolically,

$$\text{Value } (V) = \text{Results } (R) - \text{Expectations } (E)$$

Consequently, since expectations often outweigh the results, construction owners feel that they receive less value than they should. Forbes (1999) quantified the "gaps" or dissonance zones between the three parties to construction, i.e., owners, designers, and contractors in health care facilities projects. In the area of owner satisfaction factors for example, public owners and designers differed on 7 of 9 criteria, owners and contractors differed on 5 of 9 criteria, while designers and contractors disagreed on the relative importance of 2 criteria.

5. *A focus on inspection, not workmanship.* Code enforcement representatives of government agencies carry out construction inspections. Their role is to inspect critical aspects of the construction process by limited inspections on a number of items including reinforcing elements and concrete samples, but not workmanship.
6. *The growing emergence of subcontracting.* The subcontractors are often priced in a manner that does not reflect the contract with the owner—even if the owner pays a high price, the subcontractor may still have to work with inadequate budgets, often compromising quality as a result. Deming's fourth point cautions against awarding contracts based on price tags alone.
7. A culture of slow adoption of innovation—small contractors often lack the expertise or financial resources to adopt technological advances—adoption is inhibited further by fear and uncertainty. Roofing contractors, for example, tend to use the same time-honored methods to ensure that supplies and equipment are on site each day. Items that are frequently forgotten are delivered by expediters, contributing to waste in the industry.
8. The training needed often does not get to the decision-makers in the construction industry. Construction management programs around the country have been providing higher levels of training for managers; however, this training has not reached the ultimate decision-makers in the industry. Efforts to enhance quality and productivity are likely to be frustrated under this scenario.
9. Owners have not specifically demanded productivity and quality. There is a general lack of productivity/quality awareness in the industry among all parties, including owners. Owners have come to accept industry pricing—they have not been able to influence the productivity of the industry—prices have simply become higher on a per unit basis. By contrast, manufacturing activities have become cheaper over time on a per unit basis.
10. Architect/engineer (AE) contracts are said to be unclear with respect to professional standards of performance, often leading to unmet expectations. Construction owners feel that typical A/E contracts protect designers at the owner's expense. For example, prevailing contract language relieves designers of any role in the case of a lawsuit or arbitration between an owner and contractor. An outgrowth of this is the practice of "substantial completion," where a job is usable but has 5% of the remaining work in the form of a "punch list." An owner often has a very difficult time in persuading a contractor to finish that work.

11. Few large companies, and virtually no small companies have implemented the concept of a quality or productivity manager—cost-cutting trends have resulted in such a position being viewed as an unjustifiable luxury.
12. There is little, if any, benchmarking—many manufacturers and service organizations have become preeminent by adopting the best practices of benchmarked organizations. Construction has done very little of this due to distrust, fear of losing competitive advantage, but more likely, simply by being anachronistic.

6.2.5.3.1 Quality management systems. The Malcolm Baldrige Quality Award criteria provide an excellent framework for a construction organization's QM system; these criteria embody many of the concepts advocated by the quality pioneers—Deming, Juran, Crosby. Past winners of the Baldrige Award have proven to be been world-class organizations. Industrial engineers can assist construction organizations to improve quality and productivity by applying the Malcolm Baldrige criteria to their business model.

The Baldrige Award Criteria are based on a framework of core values for quality improvement comprised of seven critical areas:

1. Leadership
2. Customer and market focus
3. Strategic quality planning
4. Information and analysis
5. Human resource development
6. Process management
7. Operational results

Other industry-recognized QM systems include the ISO9000: 2000 standards.

6.2.5.3.2 Industry awards. The National Association of Home Builders created a National Housing Quality Program in 1993 to promote quality improvement in that industry. The National Housing Quality Award was developed based on the Malcolm Baldrige Award.

6.2.6 Productivity management

By definition, productivity is measured as the ratio of outputs to inputs; it may be represented as the constant-in-place value divided by inputs such as the dollar value of material and labor. In the construction environment, productivity measurements may be used to evaluate the effectiveness of using supervision, labor, equipment, materials, etc., to produce a building or structure at the lowest feasible cost.

Mali (1978) combines the terms productivity, effectiveness, and efficiency as follows:

$$\begin{aligned}\text{Productivity index} &= \frac{\text{Output obtained}}{\text{Input supplied}}\\ &= \frac{\text{Performance achieved}}{\text{Resources consumed}}\\ &= \frac{\text{Effectiveness}}{\text{Efficiency}}\end{aligned} \tag{6.1}$$

Therefore, productivity is the combination of effectiveness and efficiency. To increase productivity, the ratio(s) mentioned in Equation 6.1 must increase. This can be achieved by increasing the output, reducing the input or permitting changes in both such that the rate of increase in output is greater than that for input.

An increase in productivity can be achieved in five ways as follows:

(i) Reduced costs: $\dfrac{\text{output at same level}}{\text{input decreasing}}$

(ii) Managed growth: $\dfrac{\text{output increasing}}{\text{input increasing (slower)}}$

(iii) Reengineering: $\dfrac{\text{output increasing}}{\text{input constant}}$

(iv) Paring-down: $\dfrac{\text{output down}}{\text{input down (faster)}}$

(v) Effective working: $\dfrac{\text{output increasing}}{\text{input decreasing}}$

6.2.6.1 *Total productivity*

Total productivity (TP) is the ratio of output to all inputs. All input resources are factored in this principle. Tracking the productivity changes that occur in different time periods is the most useful application of TP. Sumanth (1984) points to the limitations of partial productivity measures, which are measured by the ratio of output to one class of input such as labor productivity. Such measures if used alone can be misleading, do not have the ability to explain overall cost increases, and tend to shift blame to the wrong areas of management control.

Total productivity may be defined as

$$\text{TP} = \frac{\text{Total sales or value of work}}{\begin{array}{c}\text{Labor cost}(M_1) + \text{Material cost}(M_2) + \text{Machinery cost}(M_3) + \text{Money cost}(M_4) \\ + \text{Management cost}(M_5) + \text{Technology cost}(M_6)\end{array}}$$

or

$$\text{TP} = \frac{T(s)}{M_1 + M_2 + M_3 + M_4 + M_5 + M_6} \tag{6.2}$$

Since

$$P_i = \frac{T(s)}{M_i}$$

$$P_t = \frac{1}{1/P_1 + 1/P_2 + 1P_3 + 1/P_4 + 1/P_5 + 1/P_6} \tag{6.3}$$

The above-mentioned factors are expressed as constant dollars (or other currency) for a reference period. To increase TP, it is necessary to determine which partial productivity factor (P_i) has the greatest short- and long-term potential effect on TP.

As pointed out by Oglesby et al. (1989), traditional construction project management tools do not address productivity; they include schedule slippages and cost overruns. Forbes and Golomski (2001) observed that the construction industry as a whole measures performance in terms of completion on time, completion within budget, and meeting construction codes.

Construction organizations (designers and constructors) would benefit significantly by establishing formal productivity and quality improvement programs that build on the knowledge gained from the measurement approaches that have been discussed above.

Industrial engineers can support such organizations in setting up productivity and quality improvement programs and providing ongoing measurement, which is critical to the process of continuous improvement.

Construction productivity is a major concern, especially when compared to other industries. As reported by the U.S. Department of Commerce, construction productivity has been rising at a much slower rate than other industries; between 1990 and 2000 it rose by approximately 0.8% compared to more than 2% for all U.S. industries. Construction costs have been increasing at the same time. Raw materials such as steel, staples have been rising, especially in the face of escalating global demand. Labor costs are a major component of most construction projects—in the vicinity of 40%, yet on many construction sites a large percentage of the daily labor hours are unproductive.

Activity sampling studies have shown that the working portion of activities generally occupies 40% to 60%, and by the same token 40% to 60% of labor hours are unproductive. There are many reasons for lost time—poor communications, waiting on assignments, waiting on resources, double material handling, rework, accidents, late or inaccurate job status reports, lack of supervision, etc. One third of these losses reflect issues that are within management's control. Construction profitability is directly linked to labor productivity. Industry-wide studies suggest that most construction projects yield net profits of 2 to 3% of the total project cost.

A hypothetical example:

Contract price	\$10,000,000
Labor cost (40%)	\$4,000,000
Other costs, overheads, etc.	\$5,700,000
Net profit	\$300,000

Assuming a 5% reduction in labor cost due to productivity improvement,

$$\text{savings in labor cost} = \$4{,}000{,}000 \times 0.05 = \$200{,}000$$

$$\text{revised net profit} = \$300{,}000 + \$200{,}000 = \$500{,}000$$

Hence, a 5% improvement in labor productivity can improve profitability by 66.7%.

Similarly, the value of lost labor hours due to management inefficiencies

$$= \$4{,}000{,}000 \times 1/3 = \$1{,}333{,}333$$