



TENTH EDITION

# COGNITION

THEORIES AND APPLICATIONS

Stephen K. Reed



## PREFACE

I began writing the first edition of *Cognition: Theory and Applications* in 1979 when I was a Visiting Associate Professor at the University of California, Berkeley. Cognitive psychology was then a relatively new field of study and I was disappointed by the lack of cognition textbooks that could inform students about this emerging field. My goal in writing a textbook was to share my own excitement by introducing students to the research, theories, and applications of cognitive psychology. Eight revisions of that first edition show its continued development.

I had not planned to write a new edition when I retired from teaching in 2014. I instead focused on writing journal articles on integrating ideas in cognitive psychology. I referred to this project as searching for the big pictures (Reed, 2020a). When I ran out of ideas for big pictures I returned to writing books (Reed, 2021) and enthusiastically began this 10<sup>th</sup> edition.

Both my new publisher (SAGE) and I approached the 10<sup>th</sup> edition as if it were a new book. It had been nine years since the publication of the 9<sup>th</sup> edition and the time in between editions provided an opportunity to introduce changes. It also produced a need to extensively update the content, and my search for the big pictures provided a head start on adding new content to the 10<sup>th</sup> edition of *Cognition*.

Another head start was my recently completed book *Cognitive Skills You Need for the 21<sup>st</sup> Century* (Reed, 2020b). The motivation for that book was the World Economic Forum's *Future of Jobs Report 2018* based on interviews of some of the world's largest employers regarding the latest needed skills and human investment trends across their industries. The report indicated that by 2022 trending skills would include active learning and learning strategies, reasoning, analytical thinking and innovation, complex problem solving, and creativity. I discuss these skills (and many others) in greater depth in this new edition of *Cognition*, making it especially relevant and useful to instructors and students in cognitive psychology courses.

You do not, however, need to wait for graduation to apply these skills. Applications are tailored to align with student interests, and they are useful in our daily lives, as shown in the photographs of people performing these activities throughout the book. The emphasis on real world applications and the visually appealing art program that brings them to life, along with the balanced coverage of classic and contemporary theories and research, reflects my attempt

to maintain the best features of previous editions while introducing new features to substantially enhance that content.

## **CHANGES IN THE TENTH EDITION**

The most visible change in the tenth edition is that it is in full color. I have enjoyed the task of finding colorful figures and photos to make this edition more engaging and informative than previous editions. As you might imagine, a delay of 10 years between the publication of the ninth and tenth editions also provided an opportunity to include extensive new material that has appeared in the past decade.

One of these changes was the addition of a new chapter on action. Many cognitive psychologists now believe that much of cognition – perception, memory, learning, language, decision making, and problem solving – occurs to support action. It now appears counterproductive to describe these cognitive processes without considering how we use them to support action. Another change was finishing the book with the chapter on ‘Expertise and Creativity’. Placing that chapter after the one on ‘Decision Making’ enables discussion of both decision making and problem solving within the context of expertise.

A frequent request by reviewers of previous editions was to include more material on cognitive neuroscience. We therefore relied on the expertise of contributing author Paul Merritt to help add that content to the book. Paul had an additional influence on content. Many of the students who enrolled in his cognitive psychology course at Georgetown University and Colorado State University were not psychology majors. I therefore attempted to make this book more user-friendly to those students by eliminating or rewriting material that appeared too difficult for a more general audience. The result is a text that undergraduate students in cognitive psychology will find very readable and accessible.

## **APPROACH**

I have continued to use three criteria for selecting material. The first is whether the material makes an important contribution to cognitive psychology. The second is whether it is accessible to students. Will they understand it and find it interesting? The third is whether it can be easily integrated with other material in the book. There must be a clear flow of ideas to tell a coherent story.

Three themes appear throughout the book: research, theories, and applications.

## 1. Research

Research is the foundation of any field of science. Cognitive psychologists have developed many innovative research paradigms to make inferences about cognitive operations. Their measures include accuracy, response times, verbal reports, brain waves, and blood flow to identify cognitive operations and discover their locations in the brain. This book describes experiments in sufficient detail to help students understand these research paradigms. The detailed descriptions typically occur early in each chapter to provide a basis for understanding the remainder of the chapter. Most chapters have a chronological flow that illustrates how contemporary research has built on the findings of classical research.

## 2. Theories

Research would consist only of a collection of measurements if theories did not organize and provide interpretations of the findings. The chapters begin with the initial theories that established a foundation for cognitive psychology such as the information-process models developed by Broadbent for speech and Sperling for vision, Miller's description of short-term memory, the Atkinson and Shiffrin model of learning, levels of processing and encoding specificity, Paivio's dual coding theory, Rosch's work on categorization, variations of semantic memory models, Kintsch's model of text comprehension, Newell and Simon's theoretical framework for problem solving, and the Kahneman and Tversky study of heuristics in decision making. The chapters then show how these theories have evolved, either as others continued to develop the initial ideas or proposed competing theories. This emphasis on the evolution of theoretical understanding is a central feature of this book.

## 3. Applications

Although research and theories are key components of science, both can appear rather abstract without illustrations of how they are useful. Previous editions of *Cognition: Theory and Applications* have always included applications but the applications received less emphasis than the theories. The tenth edition places more emphasis on applications by adding a section on applications at the end of each chapter. Identifying visual disorders, misusing

cell phones, managing cognitive overload, improving learning strategies, creating better memory codes, identifying clinical imagery, understanding dementia, producing concept maps, assessing comprehension, designing helpful decision aids, resolving conflicts, and enhancing creativity are a few of the many examples.

## ORGANIZATION

The 14 chapters in the book cover a wide range of topics, and instructors can expand on whatever topics interest them. The book is divided into three parts: Cognitive Components, Cognitive Representations, and Cognitive Skills.

Part One on Cognitive Components consists of chapters on pattern recognition, attention, working memory, long-term memory, and action. Theories of pattern recognition seek to specify how people recognize and store descriptions of patterns in memory. Theories of attention are needed to explain how we select perceptual information. Working memory enables us to combine information retrieved from long-term memory with information that arrives from the environment. But its limited capacity and fast decay rate make it necessary for us to enter into long-term memory any new information we want to remember over a long interval. Action is needed to put this information to use.

Part Two on Cognitive Representations contains chapters on memory codes, visual images, categorization, and semantic organization. The first two chapters describe qualitatively different memory codes because our ability to remember depends on the kind of memory code that we create. The study of memory codes also has important implications for how efficiently people retrieve information from memory and perform spatial reasoning tasks. The next two chapters emphasize how knowledge is organized into categories and how categories are organized into hierarchies. This organization can be studied by measuring how quickly people make classification decisions and retrieve semantic information.

Part Three on Cognitive Skills consists of chapters on language, decision making, problem solving, and expertise and creativity. Language involves not only the meaning of individual words but the combination of words to form sentences that are both grammatically correct and convey intended meanings. The study of decision making has often focused on how people combine information when evaluating alternatives. The term *risky decision making* is used to describe situations in which there is uncertainty regarding possible

outcomes. The study of problem solving describes the skills needed to solve different kinds of problems, identifies general strategies, and examines the role of memory in problem solving. The concluding chapter on expertise and creativity discusses how people use prior knowledge in reasoning and solving problems. The final section of this chapter describes recent theoretical and empirical approaches to the study of creativity.

In addition to the organization of chapters into three parts, the material in each chapter is organized into manageable sections and subsections. You should review the outline at the beginning of each chapter for an overview of the topics. You should also study the Learning Objectives at the beginning of each chapter to preview some of the major theoretical constructs that you will encounter during your reading.

## ACKNOWLEDGMENTS

I wrote the first edition of this book while spending a sabbatical year at the University of California at Berkeley. I am grateful to Case Western Reserve University and the Group in Science and Mathematics Education at Berkeley for providing financial support during that year. The Group in Science and Mathematics Education also furnished me with a stimulating environment, and the Institute of Human Learning provided an excellent library. Shortly after arriving at Berkeley, I had the good fortune to meet C. Deborah Laughton, a psychology editor at the time. She expressed confidence in the book long before it was deserved and, with the assistance of an excellent staff and first-rate reviewers, helped in the development of the text.

I am grateful to Abbie Rickard, Chelsea Neve, Jessica Miller, Lara Parra and all the others at SAGE listed on page iv who have contributed to the Tenth Edition, including [insert name of production editor, indexer, proof reader during Production]. Paul Merritt played a major role as contributing author by explaining the neuroscience of the cognitive components described in Part One. I would also like to thank the following reviewers for their helpful suggestions:

Arlo Clark-Foos, University of Michigan

Mark Cushman, William Peace University

Un So Diener, John Brown University

Doss Kanessa, Troy University

Jane Dwyer, Rivier University

Catherine Evans, Radford University  
Omar Garcia, Texas A&M University  
Patrick Garrigan, Saint Joseph's University  
John Geiger, Cameron University  
John Gomez Varon, Fordham University  
Joanna Gonsalves, Salem State University  
Ralf Greenwald, Central Washington University  
Robert J. Hines, University of Arkansas at Little Rock  
Robert Hoople, Bethune Cookman University  
Nora Isacoff, NYU, Columbia, and CUNY  
Andrew Kelly, Georgia Gwinnett College  
Meredith Lanska, CSU East Bay  
Roxane Raine, University of Memphis  
Christina Salnaitis, University of South Florida, St. Petersburg  
Gary Starr, Metropolitan State University  
Janet Trammell, Pepperdine University  
Scott Valerie, Indiana University Southeast  
Chris Wahlheim, UNC Greensboro  
Chrissy Whiting-Madison, Rogers State University  
Mark Yates, University of South Alabama  
Diana Young, Georgia College & State University  
Xiaowei Zhao, Emmanuel College

The comments of others are always welcome, and I would appreciate receiving suggestions from readers.

## ABOUT THE AUTHOR

**STEPHEN K. REED** is currently an emeritus professor of psychology at San Diego State University and a visiting scholar at the University of California, San Diego. He has also taught at Florida Atlantic University (1980–1988) and at Case Western Reserve University (1971–1980). His research on problem solving has been supported by grants from NIMH, the National Science Foundation, and the Air Force Office of Scientific Research. He is the author of numerous articles and books including *Psychological Processes in Pattern Recognition* (Academic Press, 1973), *Word Problems: Research and Curriculum Reform* (Erlbaum, 1999), *Cognitive Skills You Need for the 21<sup>st</sup> Century* (Oxford University Press, 2020) and *Thinking Visually*, second edition (Routledge, 2021).

DRAFT . NOT A FINAL PROOF .  
Do not copy, post, or distribute.



## **DEDICATION**

To the memory of my parents:

Kenneth D. Reed (1919 – 2007)

Anita M. Reed (1921 – 2012)

DRAFT. NOT A FINAL PROOF.  
Do not copy, post, or distribute.

# **BRIEF CONTENTS**

1. Introduction
2. Pattern Recognition
3. Attention
4. Short-term Working Memory
5. Long-term Memory
6. Action
7. Memory Codes
8. Visual Images
9. Categorization
10. Semantic Organization
11. Language
12. Decision Making
13. Problem Solving
14. Expertise and Creativity

DRAFT. NOT A FINAL PROOF.  
Do not copy, post, or distribute.

# DETAILED CONTENTS

## 1. Introduction

### The Growth of Cognitive Psychology

Information Processing Gathers Momentum

Higher Cognitive Processes

### Cognition's Relation to Other Fields

Cognitive Neuroscience

Artificial Intelligence

Cognitive Architectures

### Organization of This Book

SUMMARY

RECOMMENDED READING

## P A R T 1 Cognitive Components

## 2. Pattern Recognition

### Describing Patterns

Template Theories

Feature Theories

Combining Features

Structural Theories

### Information-Processing Stages

The Partial-Report Technique

Sperling's Model

### Word Recognition

The Word Superiority Effect

A Model of the Word Superiority Effect

### Scene Recognition

Goal-driven Scene Understanding

Deep Neural Networks

### Applications

Brain Pathways

Visual Disorders

SUMMARY

RECOMMENDED READING

## 3. Attention

### Attention Networks

The Alerting Network

The Orienting Network

The Executive Control Network

### Bottleneck Theories

Broadbent's Filter Model

Treisman's Attenuation Model

The Deutsch-Norman Memory Selection Model

### **Capacity Theories**

Example of a Capacity Model  
Capacity and Stage of Selection  
Multitasking

### **Automatic Processing**

When Is a Skill Automatic?  
Which Skills are Automatic?

### **Applications**

Training Drivers  
Using Cell Phones  
Coordinating Attention

SUMMARY

RECOMMENDED READING

## **4. Short-term Working Memory**

### **Forgetting**

Rate of Forgetting  
Decay versus Interference  
Release from Proactive Interference

### **Capacity**

The Magic Number 7  
Individual Differences in Chunking  
Searching Short-Term Memory

### **Working Memory**

Baddeley's Working Memory Model  
Baddeley's Revised Model  
Central Executive  
Executive Functions

### **Applications**

Working Memory and Emotion  
Instruction  
Neuroscience of Working Memory

SUMMARY

RECOMMENDED READING

## **5. Long-term Memory**

### **The Atkinson-Shiffrin Model**

Transferring Information to Long-Term Memory  
Verbal Rehearsal and Learning  
Rehearsal and the Serial Position Effect

### **Types of Memory**

Episodic Memory  
Semantic Memory  
Procedural Memory

### **Neuroscience of Long-term Memory**

Early Investigations  
Encoding and Retrieval  
Consolidation

### **Applications**

Acquisition  
Retention  
Retrieval

SUMMARY

RECOMMENDED READING

## **6. Action**

### **Action Joins Cognition**

Action Influences Perception  
Actions Can Be Simulated  
Gestures Influence Comprehension  
Gestures Influence Thinking

### **Combining Actions and Objects**

Physical Actions  
Virtual Actions  
Mental Actions

### **Organizing Actions**

Creating high-level from low-level actions  
Action Identification Theory

### **Applications**

Cognitive Offloading  
Instruction

SUMMARY

RECOMMENDED READING

## P A R T 2 Cognitive Representations

## **7. Memory Codes**

### **The Levels-of-Processing Theory**

Emphasis on Coding Strategies  
Implications for Verbal Rehearsal  
Supporting Evidence  
Structural, Phonemic, and Semantic Processing

### **Criticisms and Modifications of the Theory**

Criticisms  
Elaboration of Memory Codes  
Distinctiveness of Memory Codes

### **Encoding Specificity and Retrieval**

The Encoding Specificity Principle  
Interaction between Encoding and Retrieval Operations  
Transfer Appropriate Processing

### **Applications**

Elaborate Memory Codes  
Distinctive Memory Codes  
Transfer Appropriate Processing  
SUMMARY  
RECOMMENDED READING

## **8. Visual Images**

### **Imagery in Learning**

Memory for Pictures  
Paivio's Dual-Coding Theory  
Mnemonic Strategies

### **Imagery in Performing Cognitive Tasks**

Scanning Visual Images  
Sequential versus Parallel Processing  
Mental Transformations  
Interference

### **Evidence for Images**

Cognitive Neuroscience  
Dreams  
Limitations of Images

### **Applications**

Geographical Reasoning  
Clinical Disorders  
Training Spatial Skills

SUMMARY

RECOMMENDED READING

## **9. Categorization**

### **Concept Identification**

Discovering Rules and Attributes  
Critique of the Concept Identification Paradigm

### **Natural Categories**

Hierarchical Organization of Categories  
Loss of Categorical Knowledge  
Typicality and Family Resemblances  
Person Perception

### **Categorization Models**

Multidimensional Scaling  
Comparing Categorization Models  
Theory-Based Categorization

### **Applications**

Instruction  
Cognitive Deficits

SUMMARY

RECOMMENDED READING

## **10. Semantic Organization**

### **Verification of Semantic Statements**

- The Hierarchical Network Model
- The Feature Comparison Model
- The Spreading Activation Model

### **The Perceptual Symbols Model**

- Semantic Verification
- Autobiographical Memory
- A Neural Network Model of Semantic Integration

### **Schema Theory**

- Bartlett's Schema Theory
- Modern Schema Theory
- Scripts: Representing Sequences of Events

### **Applications**

- False Memories
- Concept Maps
- Semantic Organization in AI

SUMMARY

RECOMMENDED READING

P A R T 3 Cognitive Skills

## **11. Language**

### **Three Aspects of Language**

- Grammar (Forming Phrases)
- Meaning (Combining Words and Morphemes)
- Sound (Producing Phonemes)

### **Hierarchical Organization of Language**

- Speech Errors
- Predicting While Comprehending Language
- Temporal Stages

### **Text Comprehension**

- Local Coherence
- Global Coherence
- Memory for Text

### **Kintsch's Model of Text Comprehension**

- Processing Assumptions
- The Construction-Integration Model
- Incorporating Prior Knowledge

### **Applications**

- Comprehension Strategies and Assessment
- Persuasive Communications

SUMMARY

RECOMMENDED READING

## **12. Decision Making**

### **Making Choices**

- Compensatory Models
- Noncompensatory Models
- Selecting a Strategy
- The Adaptive Toolkit

### **Estimating Probabilities**

- Availability
- Representativeness

### **Combining Probabilities and Values**

- Expected Value
- Expected Utility
- Subjective Expected Utility

### **Risk Dimensions**

- Importance of Risk Dimensions
- Decision Frames
- Perceived Risk

### **Applications**

- Decision Aids
- Healthy Decisions
- Action-Based Decision Making

### SUMMARY

### RECOMMENDED READING

## **13. Problem Solving**

### **Classifying Problems**

- Arrangement
- Inducing Structure
- Transformation

### **Newell and Simon's Theory**

- Objectives and Method
- Theoretical Assumptions
- Means–End Analysis

### **General Strategies**

- Subgoals
- Analogy
- Representational Transfer

### **Applications**

- Design Problems
- Conflict Resolution

### SUMMARY

### RECOMMENDED READING

## **14. Expertise and Creativity**



**Expertise and Reasoning**

- Logical Reasoning
- Associative versus Analytical Reasoning
- Analogical Reasoning

**Expertise and Problem Solving**

- Search versus Implementation
- Ill-structured Problems
- Limitations of Expertise

**Creativity**

- Inventing Products Through Imagery
- Problems and Ideas
- Neuroscience of Creativity

**Applications**

- Boosts versus Nudges
- Enhancing Creativity

SUMMARY

RECOMMENDED READING

DRAFT. NOT A FINAL PROOF.  
Do not copy, post, or distribute.



Getty Images/DrAfter123

1

## INTRODUCTION

**The Growth of Cognitive Psychology**

Information Processing Gathers Momentum  
Higher Cognitive Processes

**Cognition's Relation to Other Fields**

Cognitive Neuroscience  
Artificial Intelligence  
Cognitive Architectures

**SUMMARY****RECOMMENDED READING****LEARNING OBJECTIVES**

1. Explain the growth of cognitive psychology as a field of study, including the differences between behaviorism, the information-processing perspective and the early development of artificial intelligence.
2. Discuss the benefits of close cooperation between the fields of cognitive psychology, cognitive neuroscience, and artificial intelligence.

Cognition can be defined simply as the study of the mental operations that support people's acquisition and use of knowledge. Both the acquisition and the use of knowledge involve a variety of mental skills. If you glanced at the table of contents at the beginning of this book, you saw a list of some of these skills. Psychologists who study cognition are interested in topics such as how people recognize patterns, store information in memory, use language, solve problems, and make decisions.

The purpose of this book is to provide an overview of the field of **cognitive psychology**. The book summarizes experimental research in cognitive psychology, discusses the major theories in the field, and relates the research and theories to situations that people encounter in their daily lives—for example, reading, driving, studying, designing products, solving problems in the classroom, and making decisions.

Most students are surprised to learn how much of their everyday lives is driven by cognitive processes. One major area of interest to both students and cognitive psychologists is how to improve learning—a topic that will run throughout several chapters in the text. Another important area of interest is marketing applications of cognitive psychology. For example, manufacturers spend a great deal of time trying to make their products visually distinctive and therefore

---

**cognitive psychology** The study of the mental operations that support people's acquisition and use of knowledge

easier to find on a store shelf—a direct extension of research conducted in *visual attention*. Further business applications include manipulating pricing and discounts to influence *decision-making* in purchasing—in fact, a great deal of economics is based on decision-making research.

A great deal of law and public policy has also been formulated directly from research in cognitive psychology. In a dramatic instance of law extending from cognitive research, the New Jersey Supreme Court set down a decision outlining the only legal ways in which police officers can interrogate eyewitnesses that is based entirely on research regarding the fallibility of memory (*State v. Henderson*, 2011). In the 21<sup>st</sup> century, cognitive psychology research has had other wide-ranging impacts, from the operation of airport screening checkpoints to the safe operation of subway systems to the design of roadways and sidewalks. A section on the applications of cognitive psychology occurs at the end of every chapter.

In addition to the important ways in which cognition influences our everyday lives, how we live our lives has equally important implications for our cognitive functioning. We are learning more every year about how our diet, exercise, and sleep influence critical cognitive functions such as memory and our risk for cognitive decline as we grow older. Stress, anxiety, and depression are all influenced by our lifestyles, and these in turn can have negative effects on cognition. Similarly, both prescribed and recreational drugs can influence cognition in a variety of ways that are both positive and negative.

Cognitive skills are needed in a wide variety of professions that I discuss in my book *Cognitive Skills You Need for the 21<sup>st</sup> Century* (Reed, 2020b). The book begins with the World Economic Forum's *Future of Jobs Report 2018* that asked executives at some of the world's largest employers to report on the latest skills and human investment trends in their industries ([www.weforum.org](http://www.weforum.org)). The industries include advanced materials and biotechnology, consumer and financial services, healthcare, information and communication technologies, infrastructure and urban planning, mining and minerals, transportation, travel and tourism, and professional services. Skills that will be trending in 2022 include analytical thinking and innovation, learning strategies, creativity and originality, critical thinking and analysis, reasoning, and complex problem solving. I analyze these skills in greater depth in this text. Before delving into these topics, let's take a brief look at the history of cognitive psychology.

## THE GROWTH OF COGNITIVE PSYCHOLOGY

---

It is difficult to pinpoint the exact beginning of any field of study, and cognitive psychologists would likely offer a wide variety of dates if asked when cognitive psychology began. James's *Principles of Psychology*, published in 1890, included chapters on attention, memory, imagery, and reasoning. Kohler's *The Mentality of Apes* (1925) investigated processes that occur in complex thinking. He and other Gestalt psychologists emphasized structural understanding—the ability to understand how all the parts of a problem fit together (the Gestalt). Bartlett's book *Remembering: A Study in Experimental and Social Psychology* (1932) contained a theory of memory for stories consistent with current views. There are some other important articles or books that seemed modern but did not cause a major shift in the way cognitive psychology is currently studied.

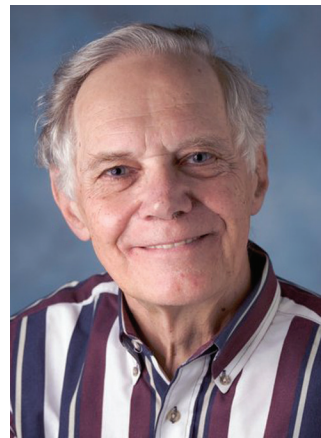
One book that had a major impact on psychological research was Watson's *Behaviorism* (1924). The book's central theme was that psychologists should become more objective by studying only what they could directly observe in a person's behavior. Watson's argument lent support to a **stimulus-response (S-R)** approach, in which experimenters record how people respond to stimuli without attempting to discover the thought processes that cause the response. The S-R approach is consistent with Watson's view because the stimulus and the response are both observable. Watson's book contributed to basing psychology on a more objective foundation of scientific observations. A limitation of the S-R approach, however, is that it does not reveal what the person does with the information presented in the stimulus.

By contrast, the information-processing approach seeks to identify how a person transforms information between the stimulus and the response. The acquisition, storage, retrieval, and use of information comprise separate stages, and the information-processing approach attempts to identify what happens during these stages (Haber, 1969). Finding out what occurs during each stage is particularly important when a person has difficulty performing a task because the psychologist can then try to identify which stage is the primary source of the difficulty. Information-processing models continue to have a major impact on our understanding of cognitive processes (Jarecki et al., 2020).

### Information Processing Gathers Momentum

Changing allegiance from a behavioral to a cognitive perspective required taking risks, as Miller (2003) points out in his personal account of the early years of the cognitive revolution. Miller (1951) wrote in the preface to his own book on language (*Language and Communication*) that the bias of the book was behavioristic. In 1951, he still hoped to gain scientific respectability by swearing allegiance to behaviorism. His later dissatisfaction with behaviorism resulted in the 1960 creation, with Jerome Bruner, of the Center for Cognitive Studies at Harvard. The cognitive emphasis at the center reopened communication with distinguished psychologists abroad, such as Sir Frederic Bartlett in Cambridge, England; Jean Piaget in Geneva, Switzerland; and A. R. Luria in Moscow, Russia. None of these three had been influenced by the behaviorism movement in the United States and therefore provided inspiration for the cognitive revolution.

Ulric Neisser's 1967 book *Cognitive Psychology* provided a clear explanation of the information-processing perspective. He defined cognitive psychology as referring "to all processes by which the sensory input is transformed, reduced, elaborated, stored, recovered, and used." This definition has several important implications. The reference



p01\_01

**PHOTO 1.1** Ulric Neisser's 1967 book *Cognitive Psychology* contributed to establishing the field of cognitive psychology.

Credit: Cornell University

**stimulus-response (S-R)** The approach that emphasizes the association between a stimulus and a response, without identifying the mental operations that produced the response

to a sensory input implies that cognition begins with our contact with the external world. Transformation of the sensory input means that our representation of the world is not just a passive registration of our physical surroundings but an active construction that can involve both reduction and elaboration. Reduction occurs when information is lost. That is, we can attend to only a small part of the physical stimulation that surrounds us, and only a small part of what we attend to can be remembered. Elaboration occurs when we add to the sensory input. For example, when you meet a friend, you may recall many shared experiences.

The storage and the recovery of information are what we call memory. The distinction between storage and recovery implies that the storage of information does not guarantee recovery. A good example of this distinction is the “tip of the tongue” phenomenon. Sometimes we can almost but not quite retrieve a word to express a particular thought or meaning. Our later recall of the word proves the earlier failure was one of retrieval rather than one of storage. The word was stored in memory; it was simply difficult to get it back out. The last part of Neisser’s definition is perhaps the most important. After information has been perceived, stored, and recovered, it must be put to good use—for example, to make decisions or to solve problems.

Neisser’s *Cognitive Psychology* (1967) brought many of these ideas together into a single source; other books on cognition followed as cognitive research and theories began to gather momentum in the 1970s. For instance, research on categorization in the 1960s had focused on a concept identification paradigm in which categories were defined by logical rules, such as category members consist of all geometric forms that are either circles or large.

The predominant criticism of the concept identification paradigm was that real-world categories, such as clothes, tools, and vehicles, are unlike the categories studied in the laboratory. A dramatic change in how psychologists viewed real-world categories had to wait until the 1970s when Eleanor Rosch (Photo 1.2) and her students at the University of California, Berkeley, began to study the characteristics of real-world categories (Rosch, 1973). Her ideas and research were so important that they deserve the extensive coverage they receive in Chapter 9 on Categorization.



**PHOTO 1.2** Eleanor Rosch made major contributions to our understanding of the organization of real-world categories.

Source: Photo of Eleanor Rosch available at [https://en.wikipedia.org/wiki/Eleanor\\_Rosch#/media/File:Eleanor\\_Rosch.jpg](https://en.wikipedia.org/wiki/Eleanor_Rosch#/media/File:Eleanor_Rosch.jpg), licensed by CC0 1.0 Universal (CC0 1.0) Public Domain Dedication

Cognitive psychology currently has widespread appeal among psychologists. Almost all psychologists studying perception, attention, learning, memory, language, reasoning, problem solving, and decision-making refer to themselves as cognitive psychologists, even though the methodology and theories vary widely across these topics. A caveat is that most of the initial contributions to the cognitive revolution were made by men because of the lack of women psychologists in academia during the time this revolution occurred. There are important exceptions, such as Eleanor Gibson's (1969) contributions to perception discussed in Chapter 2 on Pattern Recognition and Anne Treisman's (1960) theory discussed in Chapter 3 on Attention (Photo 1.3).

As a whole, the psychology field has also historically lacked racial and ethnic diversity. American Psychological Association data from 2015 showed 86% of psychologists in the U.S. workforce were white. In contrast, only 14% were from other racial or ethnic groups. Psychology is becoming more diverse as more racial and ethnic minorities enter the field, and cognitive psychology in particular is diversifying as it becomes more international, but more progress needs to be made.



**PHOTO 1.3** Anne Treisman's theory of attention advanced the information-processing approach for studying cognitive processes. Here, she is pictured receiving the National Medal of Science from President Obama.

UPI / Alamy Stock Photo

### Higher Cognitive Processes

The information-processing analysis of perception and memory was accompanied in the late 1950s by a new approach to more complex tasks. The development of digital computers

after World War II led to active work in **artificial intelligence**, a field that attempts to program computers to perform intelligent tasks, such as playing chess and constructing derivations in logic (Hogan, 1997). A seminar held at the RAND Corporation in the summer of 1958 aimed at showing social scientists how computer-simulation techniques could be applied to create models of human behavior. The RAND seminar had a major impact on integrating the work on computer simulation with other work on **human information processing**.

One consequence of the RAND seminar was its influence on three psychologists who spent the 1958–1959 academic year at the Center for Advanced Study in the Behavioral Sciences at Stanford University. The three—George Miller, Eugene Galanter, and Karl Pribram—shared a common dissatisfaction with the then-predominant theoretical approach to psychology, which viewed human beings as bundles of S-R reflexes. Miller brought with him a large amount of material from the RAND seminar, and this material—along with other recent work in artificial intelligence, psychology, and linguistics—helped shape the view expressed in their book, *Plans and the Structure of Behavior* (Miller et al., 1960).

The authors argue that much of human behavior is planned. A **plan**, according to their formulation, consists of a list of instructions that can control the order in which a sequence of operations is to be performed. A plan is essentially the same as a program for a computer. Because the authors found it difficult to construct plans from S-R units, they proposed a new unit called TOTE, an abbreviation for Test-Operate-Test-Exit. A plan consists of a hierarchy of TOTE units. Consider a very simple plan for hammering a nail into a board. The goal is to make the head of the nail flush with the board. At the top of the hierarchy is a test to determine whether the goal has been accomplished. If the nail is flush, one can exit. If the nail sticks up, it is necessary to test the position of the hammer to determine which of two operations, lifting or striking, should be performed.

The ideas expressed by Miller, Galanter, and Pribram were influenced by earlier work in two areas outside psychology. The work of Newell et al. (1958a) in the area of artificial intelligence identified strategies that people use to perform complex tasks such as playing chess. A second major influence came from linguist Noam Chomsky, who argued that an S-R theory of language learning could not account for how people learn to comprehend and generate sentences (Chomsky, 1957). His alternative proposal—that people learn a system of rules (a grammar)—was consistent with Miller, Galanter, and Pribram's emphasis on planning.

---

**artificial intelligence** The study of how to produce computer programs that can perform intellectually demanding tasks

**human information processing** The psychological approach that attempts to identify what occurs during the various stages (attention, perception, short-term working memory) of processing the stimulus

**plan** A temporally ordered sequence of operations for carrying out some task

**cognitive science** The multidisciplinary study of cognition through such fields as psychology, philosophy, artificial intelligence, neuroscience, linguistics, and anthropology



## COGNITION'S RELATION TO OTHER FIELDS

Cognitive psychology is part of a broader field of study labeled cognitive science. **Cognitive science** is the study of intelligence in humans, computer programs, and abstract theories, with an emphasis on intelligent behavior as computation (Simon & Kaplan, 1989). It attempts to unify views of thought developed by studies in psychology, linguistics, anthropology, philosophy, artificial intelligence, and the neurosciences (Hunt et al., 1989).

There are several landmarks in the development of the field (Nunez et al., 2019). The journal *Cognitive Science* began publication in 1977, and the Cognitive Science Society was formed in 1979. In 1986, the first PhD-granting cognitive science department was created at the University of California, San Diego. Don Norman played a major role in the creation of the journal, the Cognitive Science Society, and the first PhD-granting cognitive science department.

Nunez and his coauthors nonetheless argue that the field as a whole has lost impetus, focus, and recognition. Instead of becoming an interdisciplinary field, it has become a multidisciplinary field. A multidisciplinary field makes use of the theoretical perspectives of the different



**PHOTO 1.4** Don Norman at the University of California, San Diego, helped establish the multidisciplinary field of cognitive science.

Don Norman.

p01\_04

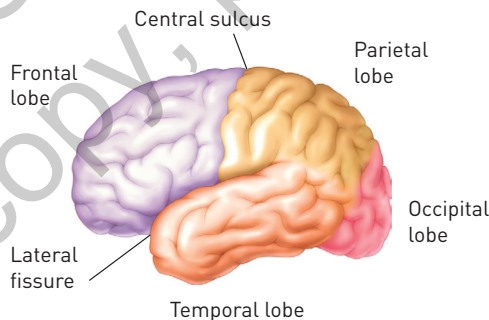
disciplines without integration. In contrast, interdisciplinary theories are more coherent and integrated. A multidisciplinary field refers to a collection of disciplines without cohesive interaction among them (Nunez et al., 2019).

Gentner (2019) agrees that the field has not converged on unified theories but argues that this is not a departure from the goals of its founders. She applauds the multidisciplinary approach to cognitive science and believes that it should be preserved and celebrated. Both the multidisciplinary and interdisciplinary approaches have made valuable contributions to cognitive science, as documented in the remainder of this chapter.

## Cognitive Neuroscience

An exception to a lack of integration among fields within cognitive science is the field of cognitive neuroscience, which combines the methodology and theories of cognitive psychology with the methods of neuroscience. Throughout the text, we will examine the relationship between specific brain areas and cognitive functions. Much of this will focus on the **neocortex**, which consists of the four lobes of the brain shown in Figure 1.1. Processing of visual information occurs in the *occipital lobe*, which is the sole function located in this brain region. The *parietal lobe* is specialized for dealing with the body and spatial information (where things are in the world, including the body). Damage to this area can result in difficulty with movement as well as loss of attention. The *temporal lobe* is essential for understanding language and contributes to recognizing complex visual patterns, such as faces. The *frontal lobe* receives sensations from all the sensory systems and contributes to planning motor movements. Damage to this area can also interfere with memory.

**FIGURE 1.1** ■ Some Major Subdivisions of the Left Hemisphere of the Neocortex.

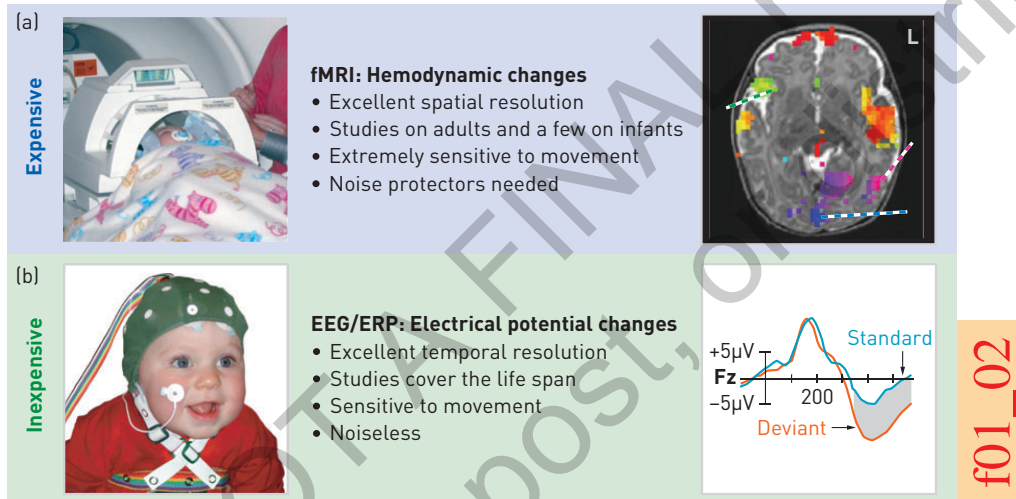


Source: Garrett, *Brain and Behavior* 5e, Figure 3.8: Lobes and Functional Areas on the Surface of the

**neocortex** Layers of the cerebral cortex that are involved in higher-order brain functions, such as perception, cognition, motor commands, and language

As technology has advanced, the ability of scientists to measure brain activity has also advanced. **Positron emission tomography (PET)** uses radioactive tracers to study brain activity by measuring the amount of blood flow in different parts of the brain (Posner & Rothbart, 1994). A more recent and widely applied method, **functional magnetic resonance imaging (fMRI)**, uses magnetic fields to measure blood flow (Figure 1.2a). It is a popular method of neuroimaging in adults because it provides high spatial-resolution maps of neural activity across the entire brain. However, the loud noises and sensitivity to movement limit its use with infants (Kuhl & Rivera-Gaxiola, 2008).

**FIGURE 1.2 ■ Functional Magnetic Resonance Imaging (fMRI) and Event-related Potentials (ERP) Provide Spatial and Temporal Measures of Brain Activity.**



Source: From "Neural substrates of language acquisition," by P. K. Kuhl & M. Rivera-Gaxiola, 2008, *Annual Review of Neuroscience*, 31, 511–534. <https://www.annualreviews.org/doi/10.1146/annurev.neuro.30.051606.094321>. Figure 1 on page 514 [row 1 and 3]

A limitation of spatial imaging techniques such as PET and fMRI is that they do not provide the precise temporal information that is necessary for analyzing many cognitive processes in which fractions of a second are theoretically important. Recording electrical activity from the scalp provides temporal precision on the order of milliseconds. The use of these **event-related potentials (ERPs)** allows scientists to study the time course of mental operations. The vertical axis of the graph in Figure 1.2b displays voltage changes and the horizontal axis displays time in milliseconds (Kuhl & Rivera-Gaxiola, 2008).

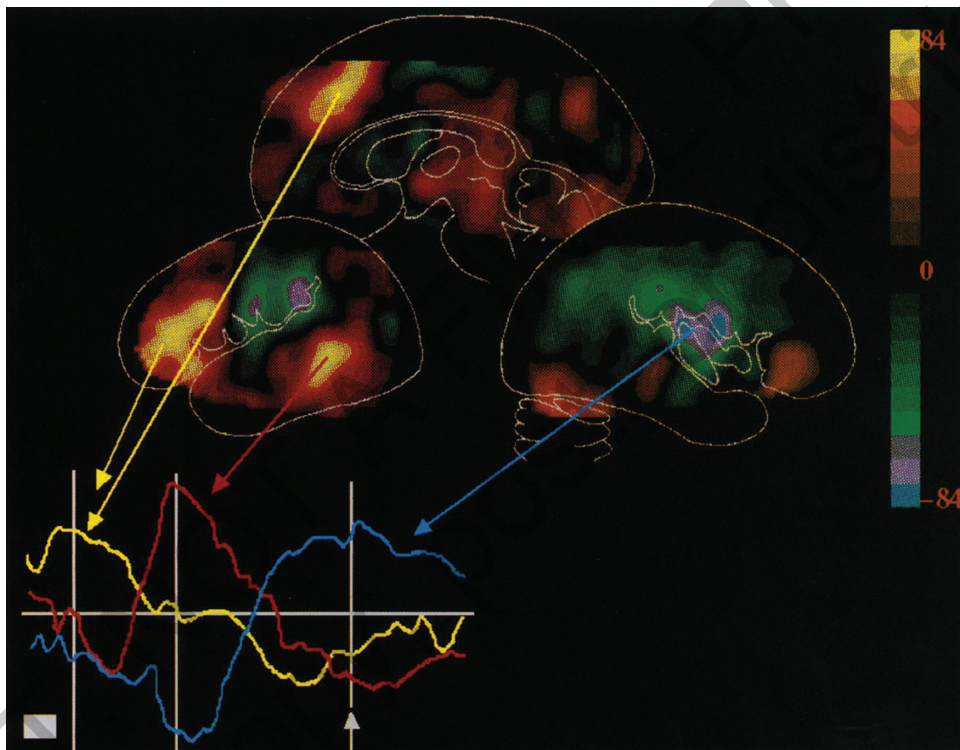
**positron-emission tomography (PET)** A diagnostic technique that uses radioactive tracers to study brain activity by measuring the amount of blood flow in different parts of the brain

**functional magnetic resonance imaging (fMRI)** A diagnostic technique that uses magnetic fields and computerized images to locate mental operations in the brain

**event-related potential (ERP)** A diagnostic technique that uses electrodes placed on the scalp to measure the duration of brain waves during mental tasks

By combining PET and ERP studies, it is possible to take advantage of the more precise spatial localization of imaging techniques and the more precise temporal resolution of electrical potentials (Posner & Rothbart, 1994). Figure 1.3 illustrates how both techniques help scientists comprehend how people understand written words (Snyder et al., 1995). The red and yellow areas show increases in blood flow, indicating that the frontal and temporal areas of the left hemisphere are important for understanding the meaning of the words.

**FIGURE 1.3** ■ A PET Scan Showing Changes in Blood Flow in the Left Hemisphere During a Cognitive Task. Brain Waves Show When Activation Occurs.

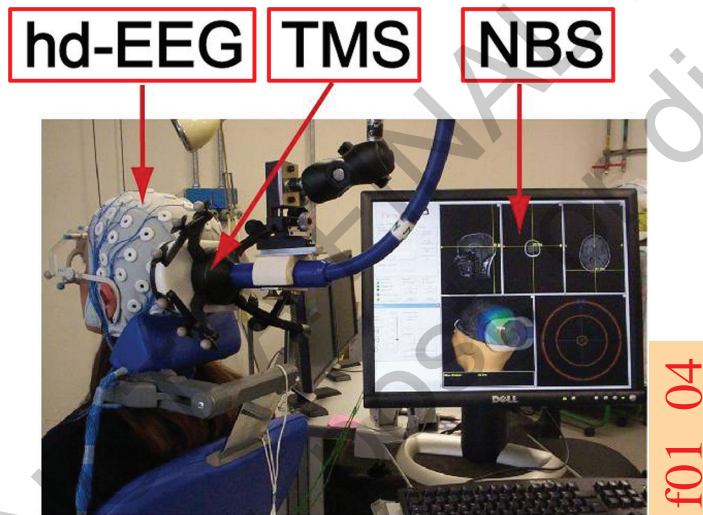


Source: From "Scalp electrical potentials reflect regional cerebral blood flow responses during the processing of written words," by A. Z. Snyder, Y. G. Abdullaev, M. I. Posner, & M. E. Raichle, 1995, *Proceedings of the National Academy of Sciences USA*, 92, 1689–1693. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC42585/>

The arrows connect PET blood-flow changes with the ERP waveforms recorded at the nearest overlying electrode on the scalp. The activation in the frontal part of the left hemisphere (yellow) leads the activation in the temporal part (red) by several hundred milliseconds. These findings imply that the earlier frontal activation is important for encoding the meaning of individual words and the later temporal activation may be more important for the integration of word meanings to understand phrases and sentences (Snyder et al., 1995). This hypothesis is consistent with the finding that damage to the temporal area of the left hemisphere often produces a language deficit that leaves the person unable to combine words to produce meaningful ideas.

Remarkably, it is possible to not only record ERPs but to directly change them (Widhalm & Rose, 2019). Transcranial magnetic stimulation produces a high-intensity magnetic field that passes through the scalp and causes neurons to fire. The effects of the stimulation are observed not only through changes in behavior but through changes in brain activity that reflect cognitive processes that contribute to that behavior (Widhalm & Rose, 2019). Figure 1.4 displays a brain stimulation system that navigates and targets **transcranial magnetic stimulation** on a 3D construction of the participant's brain. An EEG cap measures the electric field induced by transcranial stimulation and estimates its intensity on the cortical surface (Rosanova et al., 2012).

**FIGURE 1.4** ■ A Navigation Brain Stimulation System (NBS) Measures Brain Waves (EEG) Produced through Transcranial Magnetic Stimulation (TMS).



Source: From "Combining transcranial magnetic stimulation with electroencephalography to study human cortical excitability and effective connectivity," by M. Rosanova, S. Casarotto, A. Pigorinni, P. Canali, A. G. Casali, & M. Massimini, 2012, *NeuroMethods*, 67, 435-457. [https://doi.org/10.1007/7657\\_2011\\_15](https://doi.org/10.1007/7657_2011_15)

Cognitive neuroscience is particularly interesting to cognitive psychologists when it helps them evaluate cognitive theories (Yarkoni et al., 2010). For instance, we will see in Chapter 8 that one of the classic debates in cognitive psychology is the role of visual imagery in cognition. How do we know when people are using visual imagery to perform a task? Cognitive neuroscience has helped answer this question by allowing psychologists to study which part of the brain is active when

**transcranial magnetic stimulation (TMS)** A brain stimulation technique in which electrical pulses produced by a magnetic field cause neurons to fire in a focused region of the brain

people perform spatial reasoning tasks. Evidence for the use of visual imagery occurs when the same part of the brain is activated (the occipital lobe) as is activated during visual perception.

## Artificial Intelligence

In their chapter on artificial intelligence (AI) in *The Cambridge Handbook of Intelligence*, Goel and Davies (2020) propose that, from an AI perspective, the construct of intelligence is not limited to humans or even animals but includes any type of intelligent system including computers. AI implements information-processing theories that describe intelligence in terms of the content, representation, access, use, and acquisition of information. It is helpful for exploring the benefits and limitations of different ways of representing and organizing knowledge in memory. It is also helpful for exploring how robots interact with the physical world through perception and action.

There are two major paradigms for designing intelligent computers, according to the authors. Engineering AI attempts to design the smartest possible intelligent systems regardless of whether the systems reflect intelligence found in people. The vast majority of AI research on robotics and machine learning falls into this category. In contrast, psychological AI attempts to design systems that think like people.

Goel and Davies (2020) describe a paradox in which tasks that are relatively easy for computers, such as producing logical proofs and playing chess, are difficult for humans. Tasks that are relatively easy for humans, such as perceiving, walking, and talking, are difficult for computers. The goal of general AI is to make computers proficient at a wide range of tasks, including those that are easy for humans.

There are at least three benefits of close cooperation between cognitive psychologists and people working on AI (Reed, 2019). The first is that computational programs in AI can serve as potential theoretical models in cognitive psychology. An early collaborative effort between a cognitive psychologist (Alan Collins) and a computer scientist (Ross Quillian) resulted in the hierarchical network model described in Chapter 10 to represent semantic organization in human memory (Collins & Quillian 1969). But it was human problem solving (Newell & Simon, 1972) that introduced many new ideas into cognitive psychology that is described in Chapter 13.

A second benefit is that AI and cognitive psychology share common interests, such as developing methods for categorizing patterns. In his book *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*, computer scientist Pedro Domingos (2015) explains different methods used in machine learning. Cognitive psychologists have developed similar methods to evaluate models of how people categorize patterns (Reed, 2019). *The Master Algorithm* asks how these different methods can be combined to improve performance, a challenge for both AI and cognitive psychology.

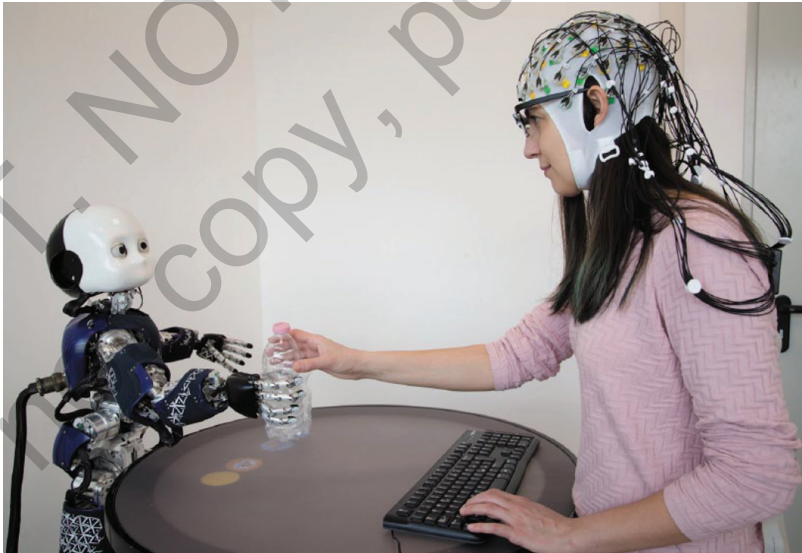
A third benefit of building bridges between AI and cognitive psychology is that the increasing impact of AI in our lives requires understanding how technology and people can work together. For instance, it is likely that robots will soon enter our lives as assistants in workplaces, shops, airports, healthcare, and classrooms (Wykowska, 2021). They will also serve as tools for



**PHOTO 1.5** Alan Newell (left) and Herb Simon (right) at Carnegie Mellon University applied concepts in artificial intelligence to model human cognition.

Getty Images/Bill Pierce

**FIGURE 1.5** ■ An Experiment that uses Behavioral Measures, Evoked Potentials, and Eye Tracking to Record the Interactions between a Participant and the Robot iCub.

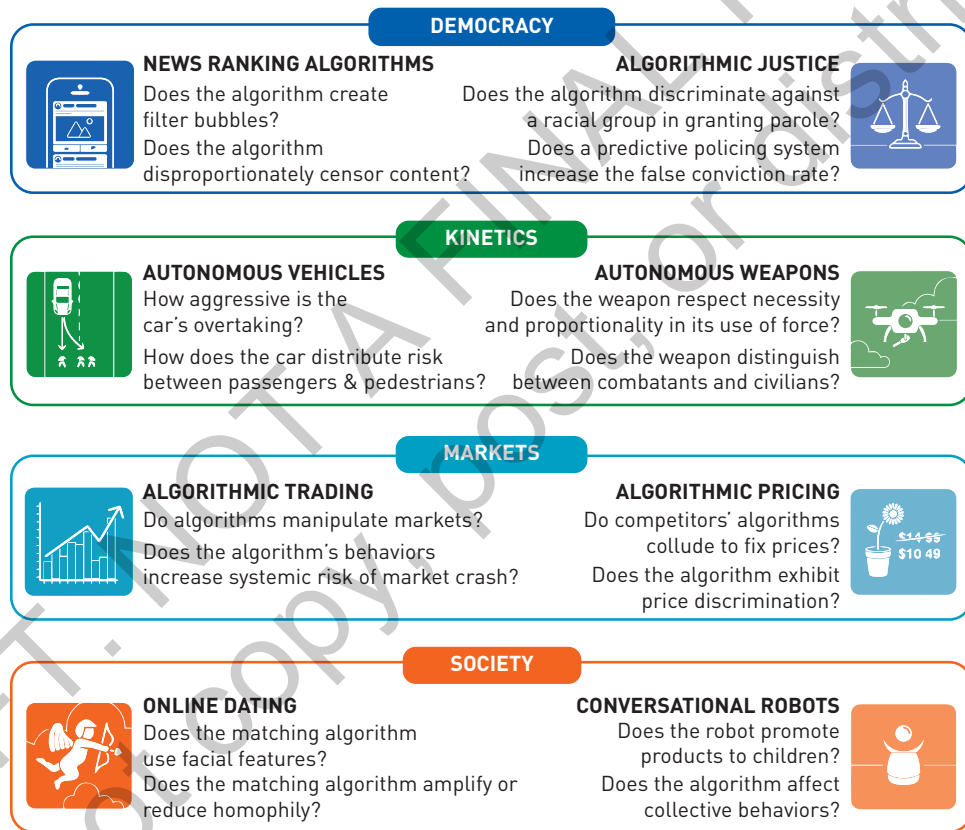


Source: From "Robots as mirrors of the human mind," by A. Wykowska, 2021, *Current Directions in Psychological Science*, 30, 34–40. doi:10.1177/0963721420978609

generating new hypotheses, predictions, and explanations regarding human cognition. Robots offer the possibility of greater experimental control over initiating and responding to interactions with people (Figure 1.5).

Although AI is already having a major positive impact on our lives, it can also have a negative impact, which has raised concern about its ethical usage. Many of the questions raised in Figure 1.6 are concerns about the social consequences of algorithms (Rahwan et al., 2019). Do they disproportionately censor content? Do they discriminate against racial groups? Do weapons use appropriate amounts of force? Do competitors collude to fix prices? These types of questions are increasingly being asked in and out of courtrooms.

**FIGURE 1.6** ■ AI Algorithms that Impact People's lives.



Source: From "Machine behavior," by I. Rahwan et al., 2019, *Nature*, 568, 477–486. <https://doi.org/10.1038/s41586-019-1138-y>

### Cognitive Architectures

A landmark in the history of artificial intelligence was a book published by Alan Newell (1990) titled *Unified Theories of Cognition*. Unified theories should ideally be able to explain all aspects

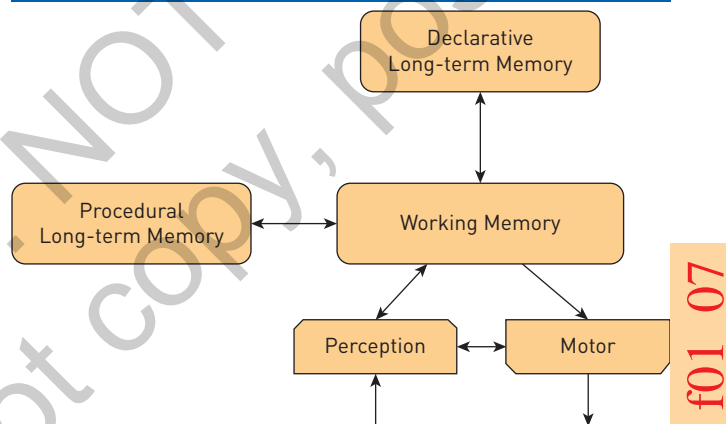


of cognition, including perception, learning, memory, problem solving, and decision-making. Such explanations require specifying interactions among the various components of cognition. Newell proposed a theory of how these components interact by developing a **cognitive architecture** called Soar (Newell, 1990).

Soar continues to be developed by the AI community (Laird, 2012), but its greatest contribution to cognitive psychology has been its influence on the development of ACT-R (Anderson, 1983)—a cognitive architecture for modeling human cognition. ACT-R assumes that cognitive architectures should be a theory of how behavior is generated through information processing that includes perception and action (Ritter et al., 2019). It has demonstrated how many aspects of cognition are intertwined, such as perception, memory, and problem solving. A manual, summer school, and workshops have supported building a community of cognitive scientists who have used the architecture to model cognition.

A limitation of ACT-R is its complexity; hence the need for summer school and workshops. Participants in a 2013 symposium on integrated cognition, sponsored by the Association for the Advancement of Artificial Intelligence, therefore met to develop a standard model of the mind based on a stripped-down cognitive architecture. A standard model would be helpful because artificial intelligence, cognitive psychology, cognitive neuroscience, and robotics all contribute to our understanding of intelligent behavior but each from a different perspective. A standard model would provide a common framework for unifying these disciplines and guide practitioners in constructing a broad range of applications.

**FIGURE 1.7** ■ The Standard Model of the Mind.



Source: From "A standard model of the mind: Toward a common computational framework across artificial intelligence, cognitive science, neuroscience, and robotics," by J. E. Laird, C. Lebiere, & P. S. Rosenbloom, 2017, *AI Magazine*, 38, 13–26. <https://doi.org/10.1609/aimag.v38i4.2744>

**cognitive architecture** An integrated system of cognitive components for modeling cognition

Figure 1.7 shows the components of the proposed standard model (Laird et al., 2017). Perception converts sensory stimuli into representations that can be stored in working memory or directly converted into actions by the motor component. Attention limits the amount of available perceptual information in both situations. Working memory provides a temporary storage space where perceptual information can be integrated with information from long-term declarative and long-term procedural memory. Declarative memory is the store for facts and concepts. Procedural memory contains knowledge about actions. The motor component uses the body to execute the actions.

You will learn much more about each of these components as you progress through the book, and the standard model will help relate these components. The authors of the standard model propose that it has the potential to provide a platform for the integration of theoretical ideas across different disciplines. I hope that including it in this text will help fulfill that goal.

## SUMMARY

One reason for studying cognitive psychology is that cognitive processes influence many aspects of our lives. It differs from behaviorism by its emphasis on mental representations and procedures, such as replacing stimulus-response associations with hierarchical plans. Cognitive psychology is a member of a multidisciplinary field labeled “cognitive science,” which also includes linguistics, anthropology, philosophy, artificial intelligence, and cognitive neuroscience. The interaction between cognitive psychology and cognitive neuroscience is the best example of an attempt to create an interdisciplinary field in which disciplines interact with each other. The objective of the standard model of the mind is to encourage more interdisciplinary interactions based on a shared cognitive architecture. The next five chapters discuss the components of this architecture—perception, working memory, long-term declarative memory, long-term procedural memory, and action.

## RECOMMENDED READING

Readers interested in how major theoretical approaches influenced the history of psychology should read Heidbreder's (1961) *Seven Psychologies*. The book contains chapters on prescientific psychology, the beginning of scientific psychology, the psychology of William James, functionalism, behaviorism, dynamic psychology, Gestalt psychology, and psychoanalysis. *The Mind's New Science: A History of the Cognitive Revolution* (Gardner, 1985) and *How the Mind Works* (Pinker, 1997) provide very readable accounts of the evolution of cognitive psychology. The article “A framework for building cognitive process models” (Jarecki et al., 2020) demonstrates how information-processing models continue to influence the field. Chipman's (2016) introduction to cognitive science appears in the *Oxford Handbook of Cognitive Science*, which she edited. *Cognitive Skills You Need for the 21<sup>st</sup> Century* (Reed, 2020b) contains 20 short chapters on themes related to this topic. *How We Learn: Why Brains Learn Better Than Any Machine... for Now* (Dehaene, 2020) an overview of many issues discussed in this textbook.



Getty Images/JurgaR

# 2

## PATTERN RECOGNITION

**Describing Patterns**

Template Theories  
 Feature Theories  
 Combining Features  
 Structural Theories

**Information-Processing Stages**

The Partial-Report Technique  
 Sperling's Model

**Word Recognition**

The Word Superiority Effect  
 A Model of the Word Superiority Effect

**Scene Recognition**

Goal-Driven Scene Understanding  
 Deep Neural Networks

**Applications**

Brain Pathways  
 Visual Disorders

**SUMMARY****RECOMMENDED READING****LEARNING OBJECTIVES**

1. Contrast feature and structural theories of pattern recognition.
2. Explain how Sperling's partial-report technique contributed to understanding characteristics of the visual sensory store.
3. Explain how the word superiority effect determines why a letter in a word is better recognized than a letter by itself.
4. Discuss the goals of understanding scenes and the applications of deep neural networks.
5. Describe how visual disorders have increased our knowledge of neural pathways.

The study of **pattern recognition** is primarily the study of how people identify the objects in their environment. Pattern recognition, which is discussed in this chapter, and attention, in the next chapter, play lead roles in the perception component of the standard model of cognition

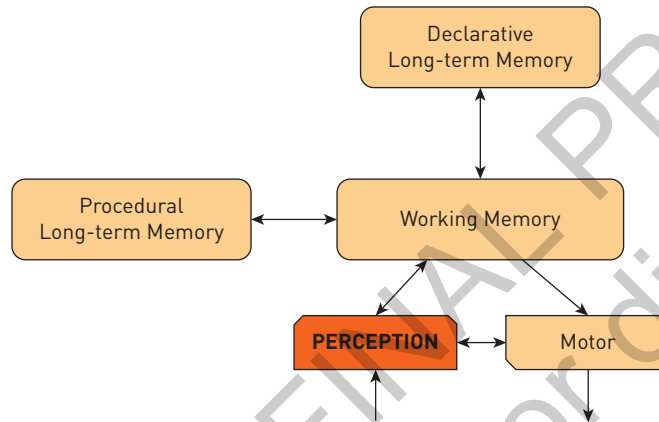
**pattern recognition** The stage of perception during which a stimulus is identified

Copyright ©2023 by SAGE Publications, Inc.

This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

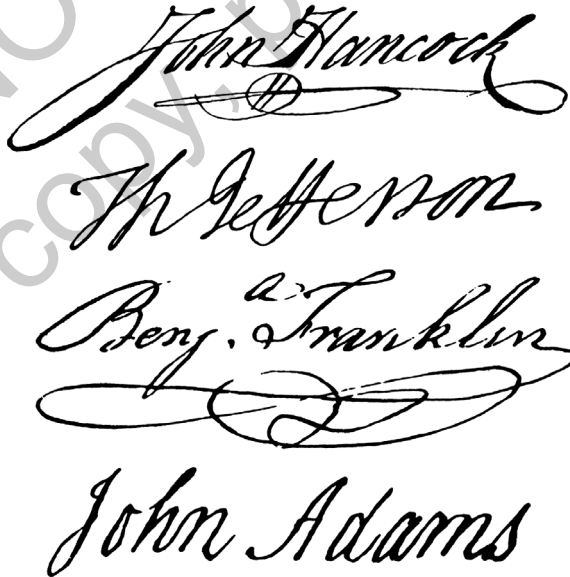
(Figure 2.1). We focus on visual pattern recognition in this chapter to provide continuity. Other chapters, such as the next chapter on attention, contain material on speech recognition.

**FIGURE 2.1** ■ The Perception Component of the Standard Model of the Mind.



Source: Based on "A standard model of the mind: Toward a common computational framework across artificial intelligence, cognitive science, neuroscience, and robotics." by J. E. Laird, C. Lebiere, & P. S. Rosenbloom, 2017, *AI Magazine*, 38, 13–26.

**FIGURE 2.2** ■ Variations in Handwriting.



Source: istockphoto.com/DNY59

Copyright ©2023 by SAGE Publications, Inc.

Our ability to recognize patterns is impressive if we stop to consider how much variation there is in different examples of the same pattern. Figure 2.2 shows various styles of handwriting. Not all people have the same style of writing, and some handwriting styles are much less legible than others. However, unless it is very illegible, we usually are successful in recognizing the words.

Our superiority over computers as pattern recognizers has the practical advantage that pattern recognition can serve as a test of whether a person or a computer program is trying to gain access to the Internet. If you have spent much time on the Internet you might have encountered a situation that required you to identify a distorted word before you were allowed to enter a site. The mangled word is easy for people to identify but difficult for computer search programs.

A large part of the literature on pattern recognition is concerned with alternative ways of describing patterns. The first section of this chapter discusses three kinds of descriptions that represent different theories of pattern recognition. The second section is about information-processing models of visual pattern recognition. The next two sections focus on word recognition and scene recognition. The last section on visual agnosia describes how studying brain disorders has contributed to establishing the neural basis of recognizing patterns.

## DESCRIBING PATTERNS

---

Consider the following explanation of how we recognize patterns. Our long-term memory (LTM) contains descriptions of many kinds of patterns. When we see or hear a pattern, we form a description of it and compare the description against the descriptions stored in our LTM. We can recognize the pattern if its description closely matches one of the descriptions stored in LTM. Although this is a plausible explanation, it is rather vague. For example, what form do these descriptions take? Let us consider three explanations that have been suggested: (1) templates, (2) features, and (3) structural descriptions.

### Template Theories

Template theories propose that patterns are really not “described” at all. Rather, **templates** are holistic, or unanalyzed, entities that we compare with other patterns by measuring how much two patterns overlap. Imagine that you made a set of letters out of cardboard. If you made a cutout to represent each letter of the alphabet and we gave you a cutout of a letter that we had made, you could measure how our letter overlapped with each of your letters—the templates. The identity of our letter would be determined by which template had the greatest amount of overlap. The same principle would apply if you replaced your cardboard letters with a visual image of each letter and used the images to make mental comparisons.

There are a number of problems with using the degree of overlap as a measure of pattern recognition. First, the comparison requires that the template is in the same position and the same orientation, and is the same size as the pattern you are trying to identify. Thus, the position,

---

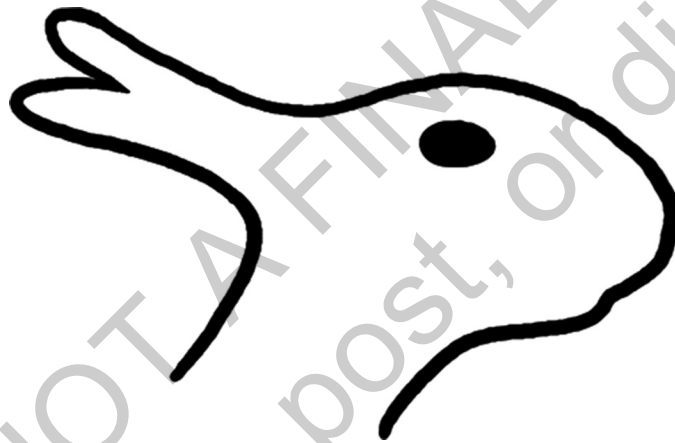
**template** An unanalyzed pattern that is matched against alternative patterns by using the degrees of overlap as a measure of similarity

orientation, and size of the templates would have to be continuously adjusted to correspond to the position, orientation, and size of each pattern you wanted to recognize. A second problem is the great variability of patterns, as was illustrated in Figure 2.2. It would be difficult to construct a template for each letter that would produce a good match with all the different varieties of that letter.

Third, a template theory doesn't reveal how two patterns differ. We could know from a template theory that the capital letters *P* and *R* are similar because one overlaps substantially with the other. But to know how the two letters differ, we have to be able to analyze or describe the letters.

A fourth problem is that a template theory does not allow for alternative descriptions of the

**FIGURE 2.3** ■ An Ambiguous Figure that can be Perceived as Either a Duck or a Rabbit.



f02\_03

Source: "What an image depicts depends on what an image means," by D. Chambers & D. Reisberg, 1985, *Cognitive Psychology*, 24, 145–174. [https://doi.org/10.1016/0010-0285\(92\)90006-N](https://doi.org/10.1016/0010-0285(92)90006-N)

same pattern. You may have seen ambiguous figures that have more than one interpretation, such as a duck or a rabbit in Figure 2.3. The two interpretations are based on different descriptions; for example, the beak of the duck is the ears of the rabbit. A template is simply an analyzed shape and so is unable to make this distinction. By contrast, a feature theory allows us to analyze patterns into their parts and to use those parts to describe the pattern.

### Feature Theories

**Feature theories** allow us to describe a pattern by listing its parts, such as describing a friend as having long blond hair, a short nose, and bushy eyebrows. Part of the evidence for feature

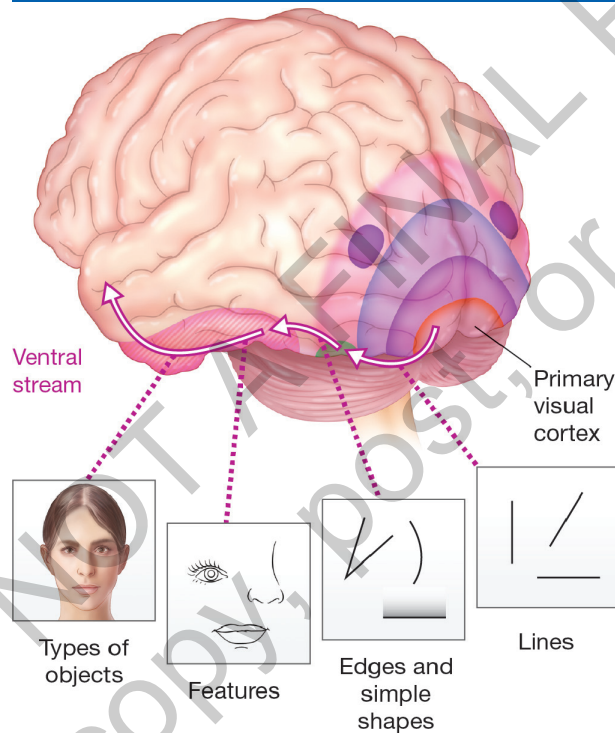
---

**feature theory** A theory of pattern recognition that describes patterns in terms of their parts, or features

theories comes from recording the action potentials of individual cells in the visual cortex. By placing microelectrodes in the visual cortex of animals, Hubel and Wiesel (1962, 1963) discovered that cells respond to only certain kinds of stimuli. Some cells might respond to a line of a certain width, oriented at a correct angle and located at the correct position in its visual field. Other cells are even concerned about the length of the line. In 1981 Hubel and Wiesel received a Nobel Prize for this work.

Figure 2.4 shows the neural processing of visual information. Light is initially detected

**FIGURE 2.4** ■ Neural Processing of Visual Features.



Source: Adapted from *INNATE: How the Wiring of Our Brains Shapes Who We Are*, by K. J. Mitchell, 2018, Princeton, NJ: Princeton University Press.

by photoreceptor cells in the retina to extract meaningful information about the visual world. This information is projected to the thalamus and areas of the primary visual cortex where the cells discovered by Hubel and Wiesel respond to features such as lines and simple shapes. These simple shapes are then combined in the ventral stream into more complex features to identify objects. We will learn more about visual features in the next section on perceptual learning and more about neural pathways in the last section on visual disorders.



## Perceptual Learning

Feature theories are convenient for explaining perceptual development, and one of the best discussions of feature theories is contained in Eleanor Gibson's (1969) *Principles of Perceptual Learning and Development*. Gibson's theory is that perceptual learning occurs through the discovery of features that *distinguish* one pattern from another.

Although most pattern recognition theorists make use of the feature concept, it is often a challenging task to find a good set of features. Gibson (1969) proposed the following criteria as a basis for selecting a set of features for uppercase letters:

1. The features should be critical ones and present in some members of the set but not in others to provide a contrast.
2. The identity of the features should remain unchanged under changes in brightness, size, and perspective.
3. The features should yield a unique pattern for each letter.
4. The number of proposed features should be reasonably small.

Gibson used these criteria, empirical data, and intuition to derive a set of features for uppercase letters. The features consist primarily of different lines and curves that are the components of letters. Examples of lines include a horizontal line, a vertical line, and diagonal lines that slant either to the right or to the left as occur in the capital letter *A*. Examples of curves include a closed circle (the letter *O*), a circle broken at the top (the letter *U*), or a circle broken at the side (the letter *C*). Most letters consist of more than one feature, such as a closed circle and a diagonal line in the letter *Q*.

A set of features is usually evaluated by determining how well it can predict **perceptual confusions**, as confusable items should have many features in common. For example, the only difference in features for the letters *P* and *R* is the presence of a diagonal line for the letter *R*; therefore, the two should be highly confusable. The letters *R* and *O* differ in many features, and so they should seldom be confused.

One method for generating perceptual confusions is to ask an observer to identify letters that are presented very rapidly (Townsend, 1971). It is often difficult to discriminate physically similar letters under these conditions, and the errors provide a measure of perceived similarity. Holbrook (1975) compared two feature models to determine how successfully each could predict the pattern of errors found by Townsend. One was the model proposed by Gibson and the other was a modification of the Gibson model proposed by Geyer and De Wald (1973). The major change in the modification was the specification of the number of features in a letter (such as two vertical lines for the letter *H*) rather than simply listing whether that feature was present.

---

**perceptual confusion** A measure of the frequency with which two patterns are mistakenly identified as each other

A comparison of the two models revealed that the feature set proposed by Geyer and De Wald was superior in predicting the confusion errors made both by adults (Townsend, 1971) and by four-year-old children (Gibson et al., 1963). The prediction of both models improved when the features were optimally weighted to allow for the fact that some features are more important than others in accounting for confusion errors. Because the straight/curved distinction is particularly important, it should be emphasized more than the others.

### Distinctive Features

Children learn to identify an object by being able to identify differences between it and other objects. For example, when first confronted with the letters *E* and *F*, the child might not be aware of how the two differ. Learning to make this discrimination depends on discovering that a low horizontal line is present in the letter *E* but not in the letter *F*. The low horizontal line is a **distinctive feature** for distinguishing between an *E* and an *F*; that is, it enables us to distinguish one pattern from the other.

Perceptual learning can be facilitated by a learning procedure that highlights distinctive features. An effective method for emphasizing a distinctive feature is to initially make it a different color from the rest of the pattern and then gradually change it back to the original color. Egeland (1975) used this procedure to teach prekindergarten children how to distinguish between the confusable letter pairs *R-P*, *Y-V*, *G-C*, *Q-O*, *M-N*, and *K-X*. One letter of each pair was presented at the top of a card with six letters below it, three of which matched the sample letter and three of which were the comparison letter. The children were asked to select those letters that exactly matched the sample letter.

One group of children received a training procedure in which the distinctive feature of the letter was initially highlighted in red—for example, the diagonal line of the *R* in the *R-P* discrimination. During the training session, the distinctive feature was gradually changed to black to match the rest of the letter. Another group of children viewed only black letters. They received feedback about which of their choices were correct, but they were not told about the distinctive features of the letters. Both groups were given two tests—one immediately after the training session and one a week later. The “distinctive features” group made significantly fewer errors on both tests, even though the features were not highlighted during the tests. They also made fewer errors during the training sessions.

Emphasizing the distinctive features produced two benefits. First, it enabled the children to learn the distinctive features so that they could continue to differentiate letters after the distinctive features were no longer highlighted. Second, it enabled them to learn the features without making many errors during the training session. The failure and frustration that many children experience in the early stages of reading (letter discrimination) can impair their interest in later classroom learning.

---

**distinctive feature** A feature present in one pattern but absent in another, aiding one's discrimination of the two patterns

**caricature** An exaggeration of distinctive features to make a pattern more distinctive

Focusing on distinctive features might aid in distinguishing among faces, as it does in distinguishing among letters. To test this, Brennan (1985) used computer-generated **caricatures** that make distinctive features even more distinctive. For instance, if a person had large ears and a small nose, the caricature would have even larger ears and an even smaller nose than the accurate drawing. When students were shown line drawings of acquaintances, they identified people faster when shown caricatures than when shown accurate line drawings (Rhodes et al., 1987). Making distinctive features more distinctive through exaggeration facilitated recognition.

### Combining Features

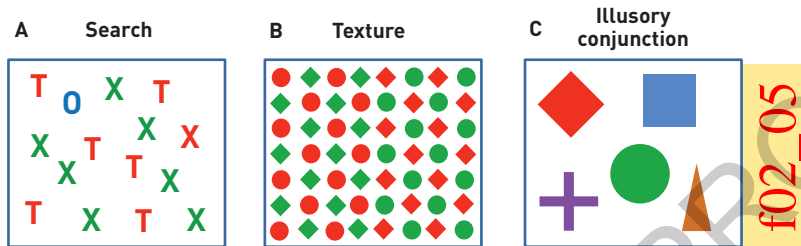
Distinctive features are a key component of our ability to locate an object in our environment. If you ever waited for your luggage at an airport, you may have noticed many people tie colorful ribbons to their luggage to help find their bags more easily because they “pop out” from the crowd. This phenomenon, as illustrated by the red flower at the beginning of the chapter, is a major prediction of feature integration theory (Treisman & Gelade, 1980).

According to this theory, all features across the entire visual landscape are represented simultaneously and pre-attentively. Thus one need only monitor the relevant feature to locate a distinctive item. Treisman and Gelade (1980) found that reaction times to find an object in a *single feature search* were independent of the size of the display, indicating that searching for a single feature is accomplished all at once. However, when two or more features must be combined in a *conjunction search*, each object in a visual scene must be examined for the combined features, which requires using attention. Returning to the airport example—if you have a standard black bag, you will now have to examine each black bag for size, shape, and so forth.

Many of the Treisman's experiments on feature integration theory explored the problem of how a perceiver combines color and shape, as these two features are analyzed by separate parts of the visual system. Figure 2.5 shows several demonstrations of how color and shape interact (Wolfe, 2018). In Panel A, it is easier to find the blue *O*, defined by the unique feature blue, than to find the red *X*. Finding the red *X* requires attending to the conjunction of red and *X* because there are also red *T*s and green *X*s in the display. Treisman found that it did not matter how many other letters were in the display, if people searched for a letter defined by a unique color or shape. The uniqueness made the letter pop out from the rest of the display, as occurs for the blue *O*. However, adding more red *T*s and green *X*s to the display would increase the time to find the red *X* because it requires attending to a conjunction of features.

Panel B illustrates another finding that is predicted by the attention requirements of feature integration theory. It is not immediately obvious that the left half of the display differs from the right half because attention is necessary for perceiving conjunctions of color and shape. The circles and diamonds switch colors, which you can observe by closely attending to the shape and color combinations. Another important implication of Treisman's theory is referred to as the “illusory conjunctions.” Following a brief glimpse of the display in Panel C, observers may report seeing an incorrect combination of color and shape, such as a

**FIGURE 2.5** ■ Visual Displays used to Evaluate Feature Integration Theory.



Source: "Ann Treisman (1935–2018)" by J. M. Wolfe, 2019, *Current Biology*, 28, R329–R341 [2018]. <https://doi.org/10.1016/j.cub.2018.03.009>

green square. Feature integration theory states that it requires attention to combine features such as color and shape. Insufficient attention, therefore, causes incorrect combinations of features.

### Structural Theories

A limitation of feature theories is that descriptions of patterns often require that we specify how the features are joined together. Describing how features join together to create a structure is a guiding principle of Gestalt psychology. To Gestalt psychologists, a pattern is more than the sum of its parts. Providing precise descriptions of the relations among pattern features was initially formalized by people working in the field of artificial intelligence who discovered that the interpretation of patterns usually depends on making explicit how the lines of a pattern are joined to other lines (Clowes, 1969).

### Structural Descriptions

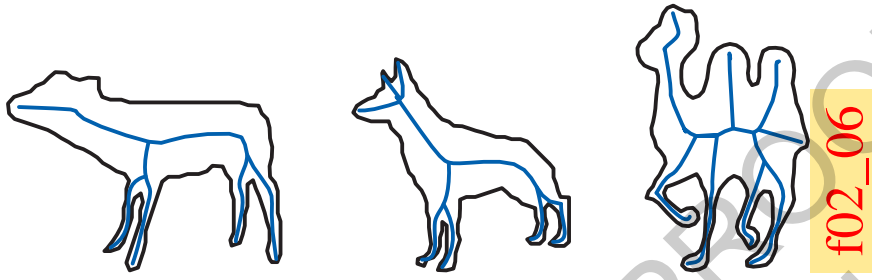
**Structural theories** describe the relations among the features by building on feature theories. Before we can specify the relation among features, we have to specify the features. A structural theory allows specification of how the features fit together. For example, the letter **H** consists of two vertical lines and a horizontal line. But we could make many different patterns from two vertical lines and a horizontal line. What is required is a precise specification of how the lines should be joined together—the letter **H** consists of two vertical lines connected at their midpoints by a horizontal line.

Figure 2.6 illustrates shape skeletons for different animals that are based on structural descriptions originally proposed by Blum (1973) as a method for distinguishing among biological forms. Wilder et al. (2011) adapted Blum's methods to make predictions about how people would classify novel shapes into categories, such as *animal* and *leaf*. Their successful predictions support the argument that people use these kinds of descriptions to make classifications. The skeleton shapes of animals have relatively curvy limbs compared to the fewer, straighter limbs of leaves.

Moving from a two-dimensional world to a three-dimensional world creates additional challenges for identifying and describing the relations among features. Figure 2.7 illustrates

**structural theory** A theory that specifies how the features of a pattern are joined to other features of the pattern

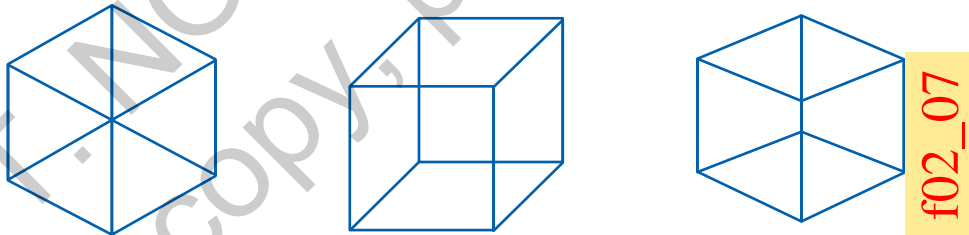
Copyright ©2023 by SAGE Publications, Inc.

**FIGURE 2.6** ■ Examples of Skeleton Structures of Animals.

Source: "Superordinate shape classification using natural shape statistics" by J. Wilder, J. Feldman, J., & M. Singh, 2011, *Cognition*, 119, 325–340. <https://doi.org/10.1016/j.cognition.2011.01.009>

the problem of identifying features by the relative difficulty of perceiving the three patterns as cubes (Kopfermann, 1930). The left pattern is the most difficult to perceive as a cube, and the pattern in the middle is the easiest. Try to guess why before reading further. (Hint: Think about the challenge of identifying features for each of the three examples.)

The theme of Hoffman's (1998) book on visual intelligence is that people follow rules in producing descriptions of patterns. The first of the many rules described in his book is to always interpret a straight line in an image as a straight line in three dimensions. Therefore, we perceive the long vertical line in the center of the right pattern in Figure 2.7 as a single line. However, it is necessary to split this line into two separate lines to form a cube because the lines belong to different surfaces. It is particularly difficult to see the figure on the left as a cube because you also need to split the two long diagonal lines into two shorter lines to avoid seeing the object as a flat pattern.

**FIGURE 2.7** ■ Perceiving Cubes.

Source: *Visual Intelligence*, by D. D. Hoffman, 1998, New York: Norton.

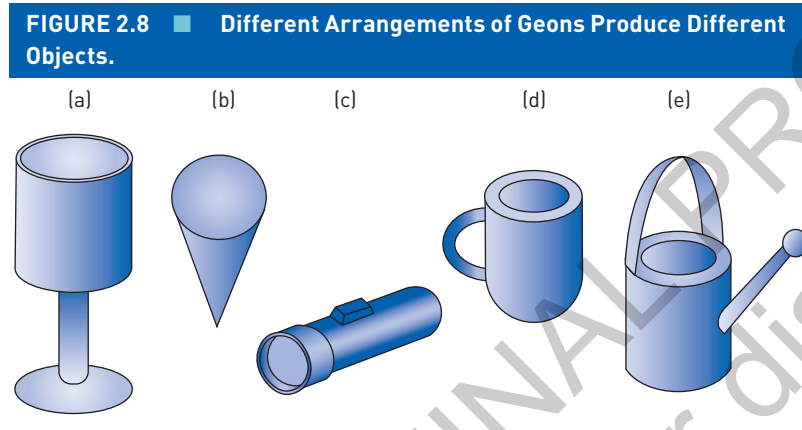
The pattern in the middle is easy to perceive as a cube, which you may have recognized as the famous Necker cube. The Necker cube is well known because your perception of the front and back surfaces of the cube changes as you view it (Long & Toppino, 2004). It is yet another example that a structural description can change when the features do not change!

### Biederman's Component Model

Descriptions of three-dimensional objects would be fairly complicated if we had to describe each of the lines and curves in the object. For example, the cubes in Figure 2.7 each consist of 12 lines

(which you may find easier to count in the left and right cubes after splitting the lines than in the reversing Necker cube). It would be easier to describe three-dimensional objects through simple volumes such as cubes, cylinders, edges, and cones than to describe all the features in these volumes.

The advantage of being able to form many different arrangements from a few components



Source: Schwartz, *Sensation and Perception 2e*: Figure 5.25

is that we may need relatively few components to describe objects. Biederman (1985) has proposed that we need only approximately 35 simple volumes (which he called **geons**) to describe the objects in the world. Some objects contain the same geons, but the geons are arranged differently. The mug (d) in Figure 2.8 would become a pail, if the handle were placed at the top rather than at the side of the container. Add two additional geons, and the pail becomes a watering can (e).

Research by Biederman et al. (2009) established that it is easier to discriminate one geon from a different geon than to discriminate two variations of the same geon. For example, U.S. college students can more easily discriminate the middle object in Figure 2.9 from the left object (a different geon with straight sides) than from the right object (a variation of the same geon with greater curvature).

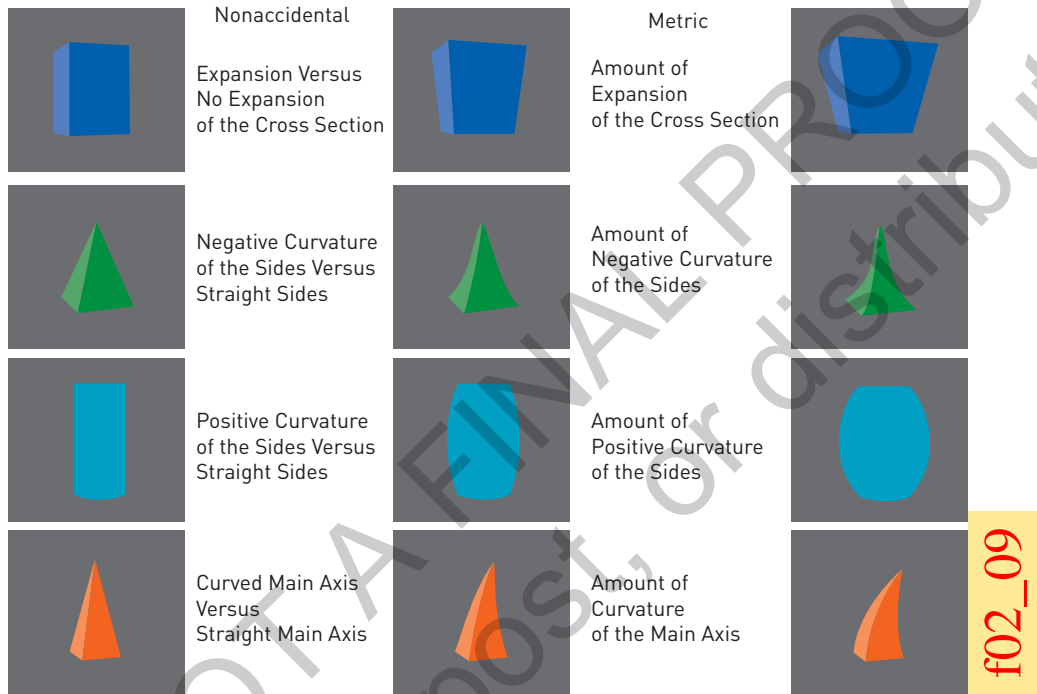
A question raised by these findings is whether there are cultural differences in people's ability to discriminate among geons. The distinction between straight lines and curves is fundamental in western culture, as we have already discovered, for discriminating among letters of the alphabet. In contrast, there is less of the need to discriminate between lines and curves by the Himba, a seminomadic people living in a remote region of Namibia. Nonetheless, the Himba also are more able to distinguish different geons from each other (the left two objects) than variations of the same geon (the right two objects).

If pattern recognition consists mainly in describing the relations among a limited set of components, then deleting information about the relations among those components should reduce people's ability to recognize patterns. To test this hypothesis, Biederman removed 65% of the contour from drawings of objects, such as the two cups shown in Figure 2.10. In the cup on the left, the contour was removed from the middles of the segments, allowing observers to see

**geons** Different three-dimensional shapes that combine to form three-dimensional objects

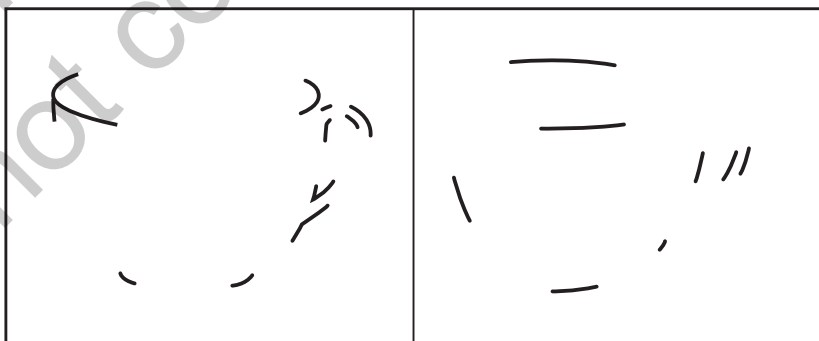
how the segments were related. In the cup on the right, the contour was removed from the vertices so observers would have more difficulty recognizing how the segments were related. When

**FIGURE 2.9** ■ Discriminating between Different Geons (Middle and Left) is Easier than Discriminating between Different Variations of the Same Geon (Middle and Right).



Source: "Representation of shape in individuals from a culture with minimal exposure to regular, simple artifacts" by I. Biederman, X. Yue, & J. Davidoff, 2009, *Psychological Science*, 20, 1437-1442. <https://doi.org/10.1111/j.1467-9280.2009.02465.x>

**FIGURE 2.10** ■ Illustration of 65% Contour Removal Centered at Either Midsegments (Left Object) or Vertices (Right Object).



Source: "Human image understanding: Recent research and theory," by I. Biederman, 1985, *Computer Vision, Graphics, and Image Processing*, 32, 29-73. [https://doi.org/10.1016/0734-189X\(85\)90002-7](https://doi.org/10.1016/0734-189X(85)90002-7)  
Copyright ©2023 by SAGE Publications, Inc.

drawings of different objects were presented for 100 msec, subjects correctly named 70% of the objects if the contours were deleted at midsegments. But if the contours were deleted at the vertices, subjects correctly named fewer than 50% of the objects (Biederman, 1985). As predicted, destroying relational information was particularly detrimental for object recognition.

In conclusion, structural theories extend feature theories by specifying how the features are related. Sutherland (1968) was one of the first to argue that if we want to account for our very impressive pattern recognition capabilities, we will need the more powerful kind of descriptive language contained in a structural theory. The experiments in this section show that Sutherland was correct. We now look at how pattern recognition occurs over time.

## INFORMATION-PROCESSING STAGES

### The Partial-Report Technique

To completely understand how people perform a pattern recognition task, we have to identify what occurs during each of the information-processing stages (pattern recognition, attention, working memory) discussed in Chapter 1. George Sperling (1960) is responsible for the initial construction of an information-processing model of performance on a visual recognition task. We discuss his experiment and theory in detail because it provides an excellent example of how the information-processing perspective has contributed to our knowledge of cognitive psychology.

Subjects in Sperling's task saw an array of letters presented for a brief period (usually 50 msec) and were asked to report all the letters they could remember from the display. Responses were highly accurate if the display contained fewer than five letters. But when the number of letters was increased, subjects never reported more than an average of 4.5 letters correctly, regardless of how many letters were in the display.

A general problem in constructing an information-processing model is to identify the cause of a performance limitation. Sperling was interested in measuring the number of letters that could be recognized during a brief exposure, but he was aware that the upper limit of 4.5 might be caused by an inability to remember more than that. In other words, subjects might have recognized most of the letters in the display but then forgot some before they could report what they had seen. Sperling, therefore, changed his procedure from a **whole-report procedure** (report all the letters) to a **partial-report procedure** (report only some of the letters).

In the most typical case, the display consisted of three rows, each containing four letters. Subjects would be unable to remember all 12 letters in a display, but they should be able to remember four letters. The partial-report procedure required that subjects report only one row. The pitch of a tone signaled which of the three rows to report: the top row for a high pitch, the middle row for a medium pitch, and the bottom row for a low pitch. The tone

---

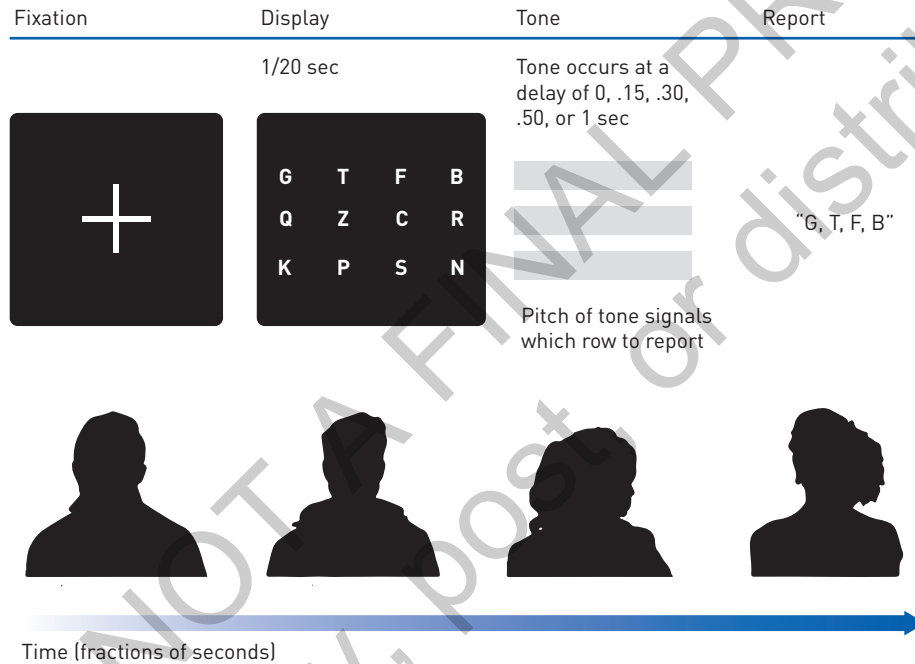
**whole-report procedure** A task that requires observers to report everything they see in a display of items

**partial-report procedure** A task in which observers are cued to report only certain items in a display of items



sounded just after the display disappeared, so that subjects would have to view the entire display and could not simply look at a single row (Figure 2.11). Use of the partial-report technique is based on the assumption that the number of letters reported from the cued row

**FIGURE 2.11** ■ Sperling's (1960) Study of Sensory Memory. After the Subjects had Fixated on the Cross, the Letters were Flashed on the Screen Just Long Enough to Create a Visual Afterimage. High, Medium, and Low Tones Signaled which Row of Letters to Report.



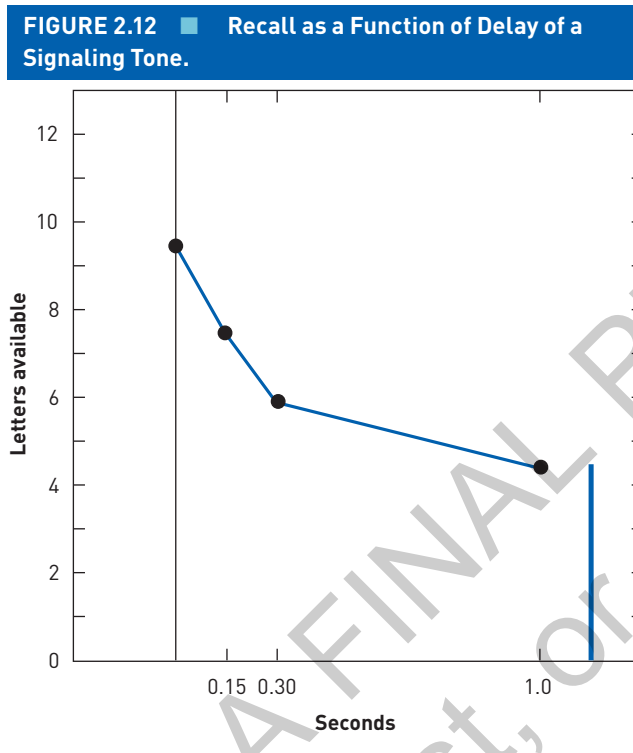
f02\_11

Photo credit: iStock/Vectorig

Source: "The information available in brief visual presentations," by G. Sperling, 1960, *Psychological Monographs*, 74 [11, Whole No. 498]. <https://doi.org/10.1037/h0093759>

equals the average number of letters perceived in each of the rows because the subjects did not know in advance which row to look at. The results of this procedure showed that subjects could correctly report three of the four letters in a row, implying that they had recognized nine letters in the entire display.

It often happens that what is best remembered about a scientist's work is not what that person originally set out to investigate. Although Sperrling designed the partial-report technique to reduce the memory requirements of his task and to obtain a "pure" measure of perception, his work is best remembered for the discovery of the importance of a visual sensory store. How did this come about? The estimate that subjects had perceived nine letters was obtained when the tone occurred immediately after the termination of the 50-ms exposure. In this case, subjects



Source: "The information available in brief visual presentations," by G. Sperling, 1960, *Psychological Monographs*, 74 (11, Whole No. 498). <https://doi.org/10.1037/h0093759>

could correctly report approximately three-quarters of the letters, and three-quarters of 12 is 9. But when the tone was delayed until one second after the display, performance declined to only 4.5 letters. That is, there was a gradual decline from nine letters to 4.5 as the delay of the tone was increased from 0 to one second (Figure 2.12).

The most interesting aspect of the number 4.5 is that it is exactly equal to the upper limit of performance on the whole-report task, as represented by the blue bar in Figure 2.12. The partial-report procedure has no advantage over the whole-report procedure, if the tone is delayed by one second or more. To explain this gradual decline in performance, Sperling proposed that the subjects were using a **visual sensory store** to recognize letters in the cued row. When they heard the tone, they selectively attended to the cued row in the store and tried to identify the letters in that row. Their success in making use of the tone depended on the clarity of information in their sensory store. When the tone occurred immediately after termination of the stimulus, the clarity was sufficient for recognizing additional letters in the cued row. But as the clarity of the sensory image faded, it became increasingly difficult to recognize additional letters. When the tone was delayed by one second, the subjects could not use the sensory store at

all to focus on the cued row, so their performance was determined by the number of letters they had recognized from the entire display that happened to be in that row. Their performance was therefore equivalent to the whole-report procedure, in which they attended to the entire display.

In 1963, Sperling proposed an information-processing model of performance on his visual report task. The model consisted of a visual information store, scanning, rehearsal, and an auditory information store. The **visual information store (VIS)** is a sensory store that preserves information for a brief period lasting from a fraction of a second to a second. The decay rate depends on such factors as the intensity, contrast, and duration of the stimulus and also on whether exposure to the stimulus is followed by a second exposure. Visual masking occurs when a second exposure, consisting of a brightly lighted field or a different set of patterns, reduces the effectiveness of the VIS.

For pattern recognition to occur, the information in the sensory store must be scanned. Sperling initially considered scanning to occur for one item at a time, as if each person had a sheet of cardboard with a hole in it just large enough for a single letter to appear.

The next two components of the model were **rehearsal** (saying the letters to oneself) and an **auditory information store** (remembering the names of the letters). To remember the items until recall, subjects usually reported rehearsing the items. Additional evidence for verbal rehearsal was found when recall errors often appeared in the form of auditory confusions—in other words, producing a letter that sounded like the correct letter. The advantage of the auditory store is that subvocalizing the names of the letters keeps them active in memory. Sperling's auditory store is part of short-term memory (STM), a topic we will consider later in the book.

Sperling revised his initial model in 1967. By this time, evidence had begun to accumulate suggesting that patterns were not scanned one at a time but were analyzed simultaneously. This distinction between performing one cognitive operation at a time (**serial processing**) and performing more than one cognitive operation at a time (**parallel processing**) is fundamental in cognitive psychology. Sperling, therefore, modified his idea of the **scan component** to allow for pattern recognition to occur simultaneously over the entire display, although the rate of recognition in a given location depended on where the subject was focusing attention.

---

**visual information store (VIS)** A sensory store that maintains visual information for approximately one-quarter of a second

**rehearsal** Repeating verbal information to keep it active in short-term memory (STM) or to transfer it into long-term memory (LTM)

**auditory information store** In Sperling's model, this store maintains verbal information in short-term memory (STM) through rehearsal

**serial processing** Carrying out one operation at a time, such as pronouncing one word at a time

**parallel processing** Carrying out more than one operation at a time, such as looking at an art exhibit and making conversation

**scan component** The attention component of Sperling's model that determines what is recognized in the visual information store (VIS)

Sperling's model was the first that attempted to indicate how various stages (sensory store, pattern recognition, and STM) combined to influence performance on a visual processing task. It contributed to the construction of information-processing models and led to the development of more detailed models of how people recognize letters in visual displays.

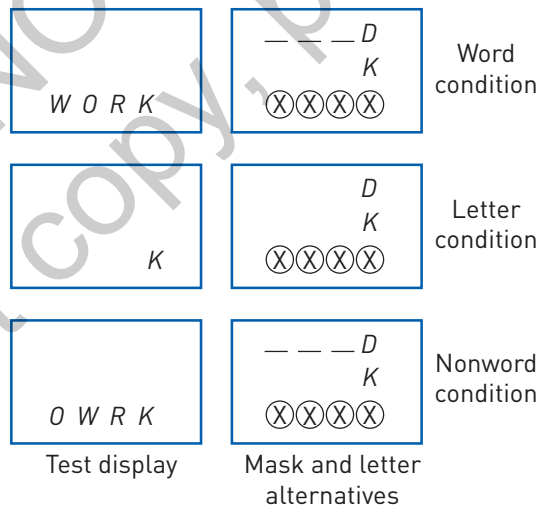
## WORD RECOGNITION

### The Word Superiority Effect

Much of the research on pattern recognition during the 1970s shifted away from how people recognize isolated letters to how people recognize letters in words. This research was stimulated by a finding that was labeled the *word superiority effect*. Reicher (1969), in his dissertation at the University of Michigan, investigated a possible implication of the scan component in Sperling's 1967 model. If the observer tries to recognize all the letters in a word simultaneously (Alderman et al., 2010), is it possible to recognize a four-letter unit in the same amount of time as it takes to recognize a single letter?

To answer this question, Reicher designed an experiment in which observers were shown a single letter, a four-letter word, or a four-letter nonword. The task was always to identify a single letter

**FIGURE 2.13** ■ Example of the Three Experimental Conditions in Reicher's (1969) Experiment. The Mask and Response Alternatives Followed the test Display. The Task was to Decide which of the Two Alternatives had Appeared in the Test Position.



f02\_13

Source: "Perceptual recognition as a function of meaningfulness of stimulus material," by G. M. Reicher, 1969, *Journal of Experimental Psychology*, 81, 275–280. <https://doi.org/10.1037/h0027768>

Copyright ©2023 by SAGE Publications, Inc.

by selecting one of two alternatives. The exposure of the stimulus was immediately followed by a visual masking field with the two response alternatives directly above the critical letter. For example, one set of stimuli consisted of the word WORK, the letter *K*, and the nonword OWRK. The two alternatives, in this case, were the letters *D* and *K*, which were displayed above the critical *K* (Figure 2.13). Observers indicated whether they thought the letter in that position had been a *D* or a *K*.

This example illustrates several characteristics of Reicher's design. First, the four-letter word has the same letters as the four-letter nonword. Second, the position of the critical letter is the same for the word and the nonword. Third, both of the response alternatives make a word (WORD or WORK) for the word condition and a nonword for the nonword condition. Fourth, the memory requirements are minimized by requiring that subjects identify only a single letter, even when four letters are presented.

The results showed that subjects were significantly more accurate in identifying the critical letter when it was part of a word than when it was part of a nonword or when it was presented alone (the **word superiority effect**). Eight of the nine subjects did better on single words than on single letters. The one subject who reversed this trend was the only subject who said that she saw the words as four separate letters, which she made into words; the other subjects said that they experienced a word as a single word, not as four letters making up a word.

The word superiority effect is an example of top-down processing. It demonstrates how our knowledge of words helps us to more rapidly recognize the letters within a word. Top-down processing, based on knowledge stored in LTM, can aid pattern recognition in different ways. Top-down processing also helps us recognize words in sentences because the sentence constrains which words can meaningfully fit into the sentence.

### A Model of the Word Superiority Effect

One of the great challenges for psychologists interested in word recognition has been to explain the reasons for the word superiority effect (Pollatsek & Rayner, 1989). A particularly influential model, the **interactive activation model** proposed by McClelland and Rumelhart (1981), contains several basic assumptions that build on the assumptions of Rumelhart's earlier model of letter recognition. The first assumption is that visual perception involves parallel processing. There are two different senses in which processing occurs in parallel. Visual processing is spatially parallel, resulting in the simultaneous processing of all four letters in a four-letter word. This assumption is consistent with Sperling's parallel scan and with Rumelhart's model of how people attempt to recognize an array of letters.

Visual processing is also parallel in the sense that recognition occurs simultaneously at three different levels of abstraction. The three levels—the feature level, the letter level, and the word level—are shown in Figure 2.14. A key assumption of the interactive activation model is that the three levels interact to determine what we perceive. Knowledge about the words of a language

---

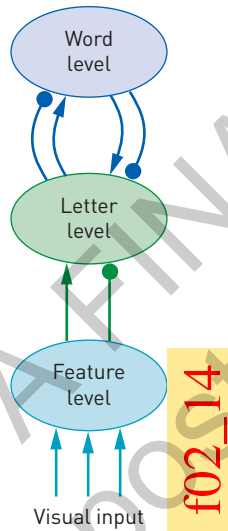
**word superiority effect** The finding that accuracy in recognizing a letter is higher when the letter is in a word than when it appears alone or is in a nonword

**interactive activation model** A theory proposing that both feature knowledge and word knowledge combine to provide information about the identity of letters in a word

interacts with incoming feature information to provide evidence about which letters are in the word. This is illustrated by the arrows in Figure 2.14, which show that the letter level receives information from both the feature level and the word level.

There are two kinds of connections between levels: excitatory connections and inhibitory connections. **Excitatory connections** provide positive evidence, and **inhibitory connections** provide negative evidence about the identity of a letter or word. For example, a diagonal line provides positive evidence for the letter *K* (and all other letters that contain a diagonal

**FIGURE 2.14** ■ The Three Levels of the Interactive Activation Model, with Arrows Indicating the Excitatory Connections and Circles Indicating Inhibitory Connections.



Source: "An interactive-activation model of context effects in letter perception: Part 1. An account of basic findings" by J. L. McClelland, & D. E. Rumelhart, 1981, *Psychological Review*, 88, 375–407. <https://doi.org/10.1037/0033-295X.88.5.375>

line) and negative evidence for the letter *D* (and all other letters that do not contain a diagonal line). Excitatory and inhibitory connections also occur between the letter level and word level, depending on whether the letter is part of the word in the appropriate position. Recognizing

**excitatory connection** A positive association between concepts that belong together, as when a diagonal line provides support for the possibility that a letter is a K

**inhibitory connection** A negative association between concepts that do not belong together, as when the presence of a diagonal line provides negative evidence that a letter is a D

**parallel distributed processing (PDP)** When information is simultaneously collected from different sources and combined to reach a decision

**neural network model** A theory in which concepts (nodes) are linked to other concepts through excitatory and inhibitory connections to approximate the behavior of neural networks in the brain

that the first letter of a word is a  $W$  increases the activation level of all words that begin with a  $W$  and decreases the activation level of all other words.

The interactive activation model was the first step for McClelland and Rumelhart in their development of neural network models of cognition. They referred to such models as **parallel distributed processing (PDP)** models because information is evaluated in parallel and is distributed throughout the network. A **neural network model** consists of several components (Rumelhart et al., 1986), some of which we have already considered in the interactive activation model. These include the following:

1. A set of processing units called **nodes**. Nodes are represented by features, letters, and words in the interactive activation model. They can acquire different levels of activation.
2. A pattern of connections among nodes. Nodes are connected to one another by excitatory and inhibitory connections that differ in strength.
3. Activation rules for the nodes. **Activation rules** specify how a node combines its excitatory and inhibitory inputs with its current state of activation.
4. A state of activation. Nodes can be activated to various degrees. We become conscious of nodes that are activated above a threshold level of conscious awareness. For instance, we become consciously aware of the letter  $K$  in the word  $WORK$  when it receives enough excitatory influences from the feature and word levels.
5. Output functions of the nodes. The output functions relate activation levels to outputs—for example, what threshold has to be exceeded for conscious awareness.
6. A learning rule. Learning generally occurs by changing the weights of the excitatory and inhibitory connections between the nodes, and the learning rule specifies how to make these changes.

The last component—the learning component—is one of the most important features of a neural network model because it enables the network to improve its performance. An example would be a network model that learns to make better discriminations among letters by increasing the weights of the distinctive features—those features that are most helpful for discriminating.

By 1992, the neural network approach had resulted in thousands of research efforts and an industry that spends several hundred million dollars annually (Schneider & Graham, 1992). The excitement of this approach can be attributed to several reasons. First, many psychologists believe that neural network models more accurately portray how the brain works than other, more serial models of behavior. Second, adjusting the excitatory and inhibitory weights that

---

**nodes** The format for representing concepts in a semantic network

**activation rule** A rule that determines how inhibitory and excitatory connections combine to determine the total activation of a concept

link nodes allows a network to learn, and this may capture how people learn. Third, the models allow for a different kind of processing in which multiple constraints (such as evidence from

both the feature and word levels) can be simultaneously considered. Neural network models have continued to be developed into one of the most powerful learning methods in AI, as indicated by their application to recognizing scenes.

## SCENE RECOGNITION

Word recognition differs from letter recognition because words are composed of interacting letters. Similarly, scene recognition differs from object recognition because scenes are composed of interacting objects that are typically arranged in a meaningful spatial layout. Recognizing objects in scenes is often driven by accomplishing goals, as explained in the next section.

### Goal-Driven Scene Understanding

Although our physical environment is usually stable, our goals can change and determine how we interact with the environment. Figure 2.15(A) illustrates four goals of scene understanding based on recognition, visual search, navigation, and action. Recognition determines whether a scene belongs to a certain category (a beach) or depicts a particular place (my living room). Visual search involves locating specific objects within the scene, such as sand, a bridge, or a lamp. Navigation determines whether it is possible to reach a particular location, such as crossing a stream. Action encompasses a broad set of activities, such as swimming, hiking, and watching television.

The four questions at the top of the figure are examples of questions we might ask for each of the different scenes (Malcolm et al., 2016). The first question “What is the scene?” requires scene recognition. It begins with gist—the perceptual and semantic information acquired from a single glance. Gist can include a conceptual understanding (a party), the spatial layout of the environment, and a few objects. It depends on the familiarity of stored representations, such as furniture is found in a living room. Unfamiliar scenes require more processing time than a brief glance to achieve scene understanding.

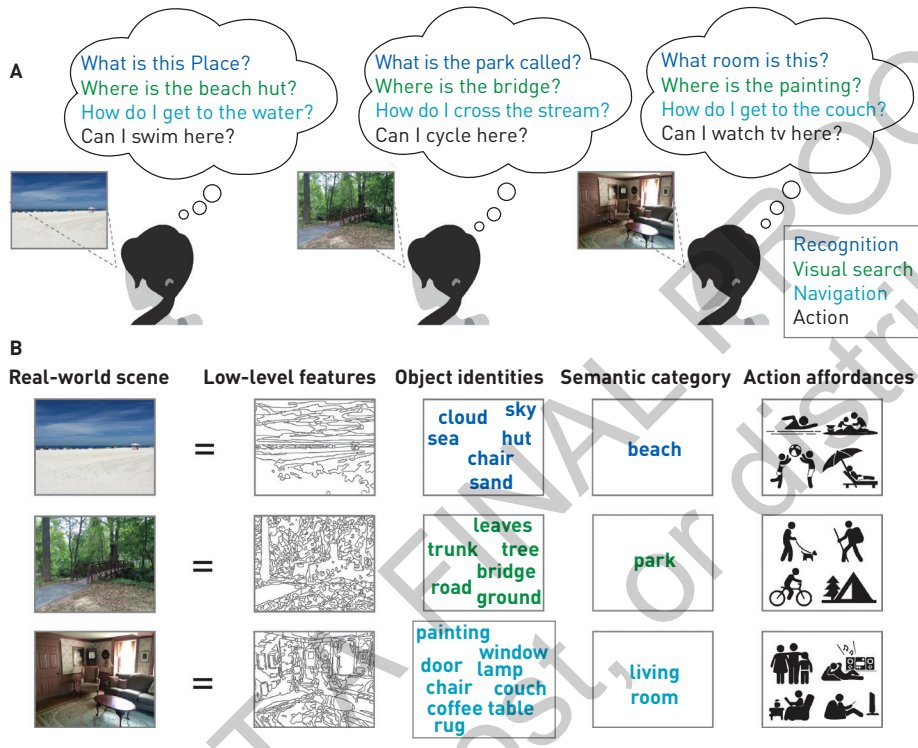
The second question “Where is X?” requires visual search using eye movements rather than a quick glance. Eye fixations focus on particular objects rather than the overall environment. They are required to answer the third question “How do I get from A to B?” Answering this question requires finding paths and potential obstacles that could block navigation, such as approaching objects. The last question “What can I do here?” determines actions, the topic of Chapter 6. Figure 2.15(B) shows scene properties that are needed to fulfill these goals. Low-level features, such as edges, establish the identity of objects. Object identities determine semantic categories and the actions that can be performed in those environments.

---

**deep neural networks** Networks that learn by adjusting thousands of connections in multiple layers

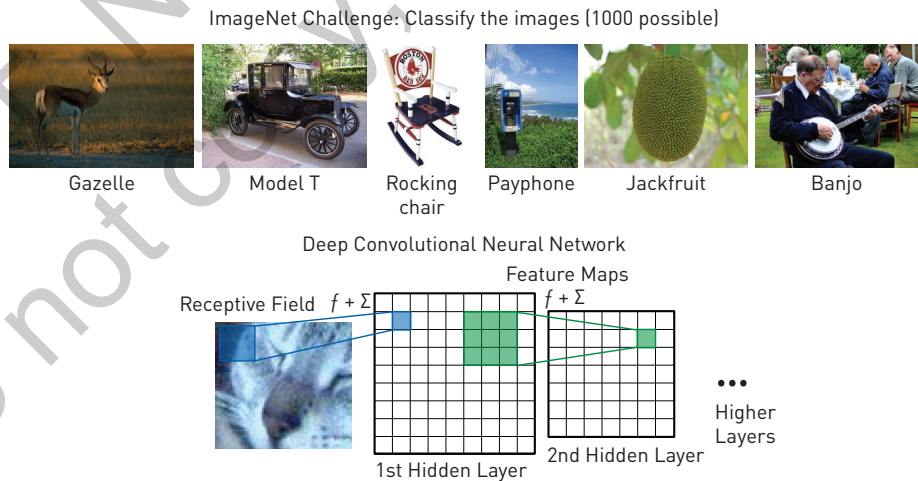


**FIGURE 2.15** ■ Goal-driven Scene Recognition.



Source: "Making sense of real-world scenes" by G. L. Malcolm, I. I. A. Groen, & C. I. Baker, 2016, *Trends in Cognitive Sciences*, 20, 843–856. <https://doi.org/10.1016/j.tics.2016.09.003>

**FIGURE 2.16** ■ Application of a Deep Neural Network to Classify Images.



Source: "Comparing the visual representations and performance of humans and deep neural networks," by R. A. Jacobs & C. J. Bates, 2019, *Current Directions in Psychological Science*, 28, 34–39.

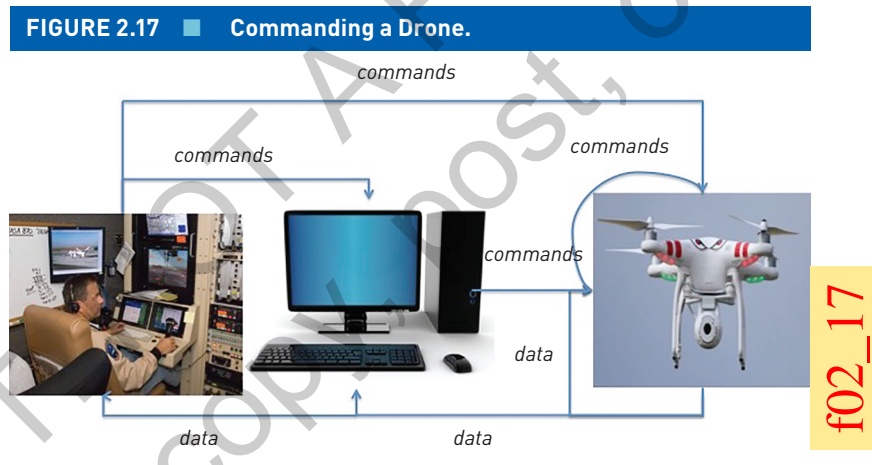
Copyright ©2023 by SAGE Publications, Inc.

## Deep Neural Networks

Computer scientists continued to develop the neural network models of the 1980s and connectionist models based on deep neural networks, which are some of the great success stories of AI (Sejnowski, 2018). **Deep neural networks** utilize the same principles as simpler networks but have added multiple layers of connections to fine-tune the weights of thousands of connections.

Figure 2.16 illustrates the application of deep neural networks to image recognition. The input begins with pixels from the image, and the output classifies the image as one of 1000 possible pictures. In between are many hidden layers in which each layer receives input from a small number of units in the previous layer to establish more global connectivity. The layers are hidden because, in contrast to the three layers in Figure 2.14, their function can be difficult to interpret.

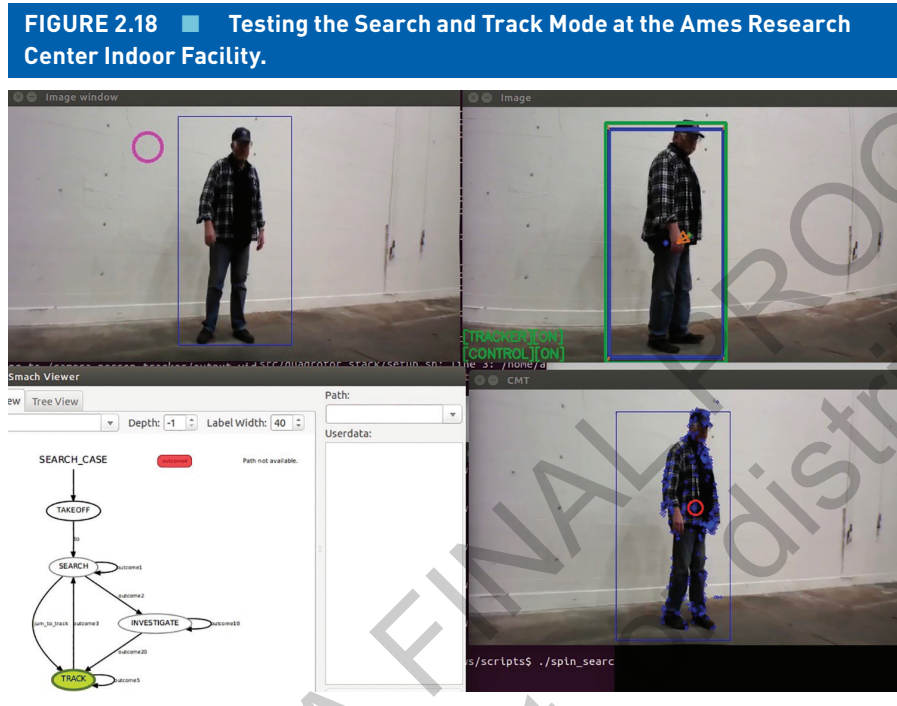
The authors of this article, Robert Jacobs and Christopher Bates at the University of Rochester's Department of Brain and Cognitive Sciences, review evidence that people are still superior at recognizing images under adverse conditions. The authors list several reasons for our perceptual advantage over machines. We learn to recognize objects in perceptually rich, dynamic, interactive environments whereas networks are trained on static images. We can take advantage of three-dimensional features whereas networks are more limited to two dimensions. A disadvantage for people, however, is our capacity limits because we cannot visually perceive



Source: "Addressing autonomy in conceptual design" by R. Morris & A. Chakrabarty, 2019, *AI Magazine*, 40, 3–16. <https://doi.org/10.1609/aimag.v40i2.2856>

and represent all aspects of a scene. These limits can nonetheless occasionally be an asset when we learn to focus on the more discriminative features.

Artificial intelligence continues to improve and the use of drones illustrates how pattern recognition can be shared by people and machines (Morris & Chakrabarty, 2019). Drones are limited by small payload capabilities and onboard processing power so some of the computational demands are offloaded to a ground computer (Figure 2.17). Joint activity between the equipment and the operator requires sensing, planning, and communication based on coordination between people and machines.



Source: "Addressing autonomy in conceptual design" by R. Morris & A. Chakrabarty, 2019, *AI Magazine*, 40, 3–16. <https://doi.org/10.1609/aimag.v40i2.2856>

Figure 2.18 shows an application to a search and track task at the Ames Research Center in California. A controller monitors the search for a target and switches to a track mode if the target is found. At some point, the target may take evasive maneuvers that require the controller to switch back to the search mode. A key component of this interaction is the boundary between human and machine decisions. For instance, humans at the console may control the search phase and allow the drone to conduct the tracking phase. Although Morris and Chakrabarty (2019) focus on searching and tracking, they argue that many of the design principles apply to integrating human decision-making with various types of devices. An important decision related to the ethical use of AI discussed in Chapter 1 requires determining who should be tracked.

## APPLICATIONS

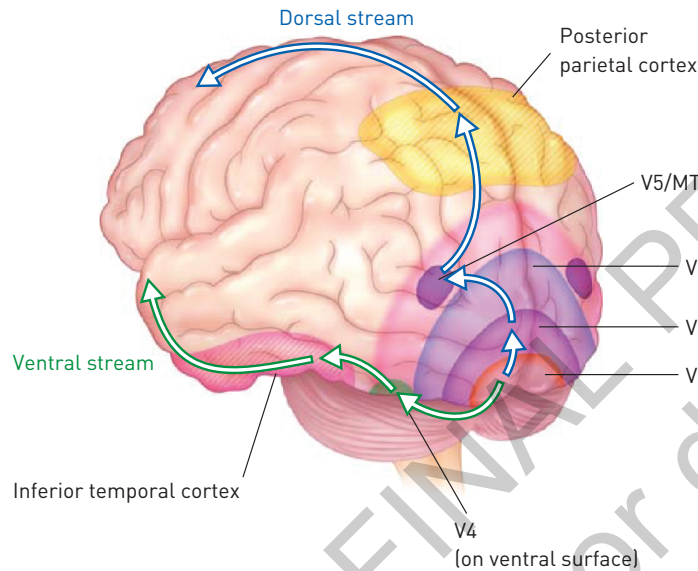
### Brain Pathways

People's remarkable ability to recognize patterns occasionally falls victim to various types of visual disorders, and studying these disorders has contributed to our understanding of visual perception (Haque et al., 2018). For example, patients with brain damage have revealed a dissociation between knowing what an object is and knowing where the object is located. Damage to one part of the brain results in an impairment of the ability to recognize visual stimuli, whereas damage to another part of the brain results in an impairment of the ability to indicate their spatial location.

Copyright ©2023 by SAGE Publications, Inc.

This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

**FIGURE 2.19** ■ Brain Pathways for Spatial (Dorsal) and Object (Ventral) Identification.



Source: Garrett, *Brain and Behavior* 6e [9781544373485], Figure 10.29: The Ventral "What" and Dorsal "Where" Streams of Visual Processing.

These impaired aspects of vision are similarly impaired in visual imagery (Levine et al., 1985). A patient with object identification difficulties was unable to draw or describe the appearance of familiar objects from memory, despite being able to draw and describe in great detail the relative locations of landmarks in his neighborhood, cities in the United States, and furniture in his hospital room. A patient with object localization difficulties could not use his memory to perform well on the spatial localization tasks but could provide detailed descriptions of the appearance of a variety of objects.

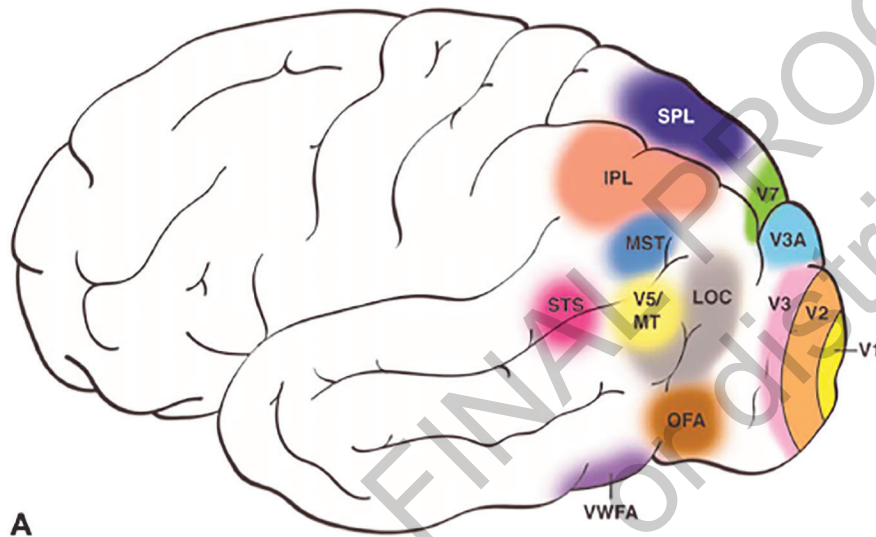
Figure 2.19 illustrates the two pathways that support the localization and the identification of objects. The *where* pathway is primarily associated with object location and spatial attention. It is often referred to as the dorsal pathway because it is located in the dorsal (or upper) part of the brain. The dorsal pathway runs upward to the parietal lobes and has strong connections with the frontal lobe that coordinates limb and eye movements.

The other pathway, which results in object recognition, is known as the *what* pathway. It travels from the primary visual cortex in the occipital lobe and processes information such as shape, size, and color, as previously illustrated in Figure 2.4. It is primarily located in the temporal lobes and is often referred to as the ventral pathway because it is located in the ventral (or lower) part of the brain.

Figure 2.20 shows the approximate locations of specialized areas for object recognition. Some parts of the brain—the occipital face area (OFA)—respond more to faces than to other types of objects. Although this area is best activated by faces, it can also be activated by other

**visual agnosia** An impairment in the recognition of visual objects

**FIGURE 2.20** ■ Specialized Areas of the Brain. Areas Discussed in the Text Include the Parietal Lobe (SPL and IPL), the Temporal Lobe (MST and MT), the Visual Cortex (V1-V7), the Occipital Face Area (OFA), and the Visual Word Form Area (VWFA).



A

f02\_20

Source: "The visual agnosias and related disorders" by S. Haque, M. S. Vaphiades, & C. J. Lueck, 2018, *Journal of Neuro-Ophthalmology*, 38, 379–392. <https://doi.org/10.1097/WNO.0000000000000556>

objects, particularly if the person has acquired previous expert knowledge about those objects (Haque et al., 2018). The visual word form area (VWFA) is activated during reading.

### Visual Disorders

Much of our knowledge of how the brain recognizes patterns comes from studies of patients with *visual agnosia*. Visual agnosia is a general disruption in the ability to recognize objects. Agnosia patients have normal visual acuity and generally show no memory deficits. The disability is also limited to a single sensory modality—for example, if you show a patient a set of keys, he will not be able to recognize them; however, if you hand him the keys to feel, he will easily identify them as keys. There are specialized forms of this disorder, such as an inability to recognize faces or familiar places.

Two general categories of agnosia disorders are **apperceptive agnosia** and associative agnosia (Farah, 2004). Apperceptive agnosia disrupts the ability of patients to group visual elements into contours, surfaces, and objects (Farah, 2004). Evidence from these patients demonstrates the pattern recognition is normally hierarchical—starting with simple cells

**apperceptive agnosia** An inability to combine visual features into contours, surfaces, and objects

**FIGURE 2.21** ■ Object Identification Tests.

**Taken from the  
*Hooper visual  
organization***



**Taken from the  
*15-object test***

f02\_21

Source: "Neuro-visual rehabilitation" by N. Raz & N. Levin, 2017, *Journal of Neurology*, 264, 1051–1058. <https://link.springer.com/article/10.1007/s00415-016-8291-0>

in the primary visual cortex and then combining these features to form a perception of the whole object, such as the face in Figure 2.4. The fusiform gyrus in the temporal lobe is of critical importance to this process (Konen et al., 2011), as is the lateral occipital cortex (Ptak et al., 2014). It is the last stage of combining features that impairs people with visual agnosia.

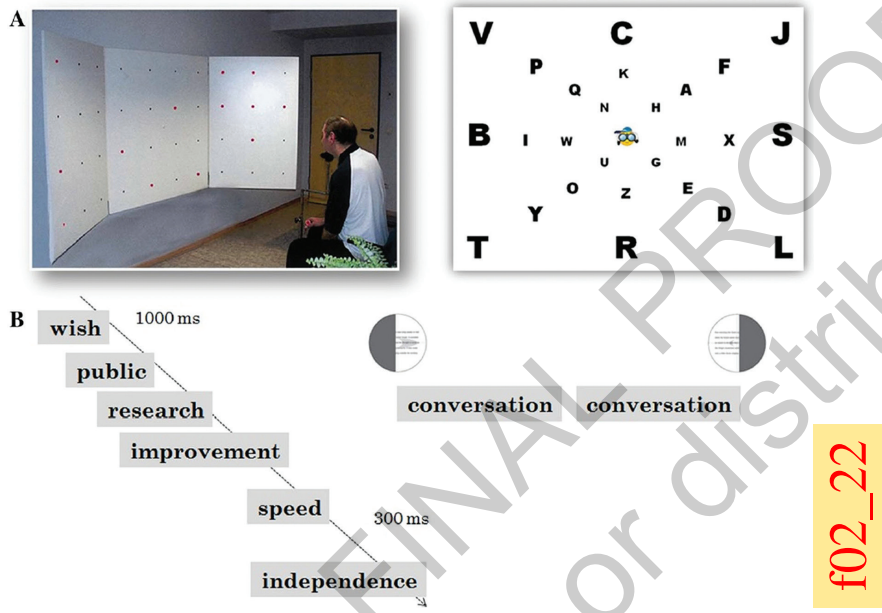
Inadequate eye movements contribute to the failure to combine visual features (Raz & Levin, 2017). A patient with apperceptive agnosia identified an object as a bird from a visual organization test, shown in the left panel of Figure 2.21. He identified the circled fragment as a beak but ignored the rest of the picture. In the right panel from an overlapping-figures test, a patient hesitated in deciding whether the circled fragment was an arrow or the ear of a cat. His eye movements did not track the length of the object to determine its identity.

Inadequate eye movements also occur in reading when patients perform shorter and delayed eye movements that limit their ability to integrate letters (Raz & Levin, 2017). A training task,

**associative agnosia** An ability to combine visual features into a whole object but an inability to recognize that object

Copyright ©2023 by SAGE Publications, Inc.

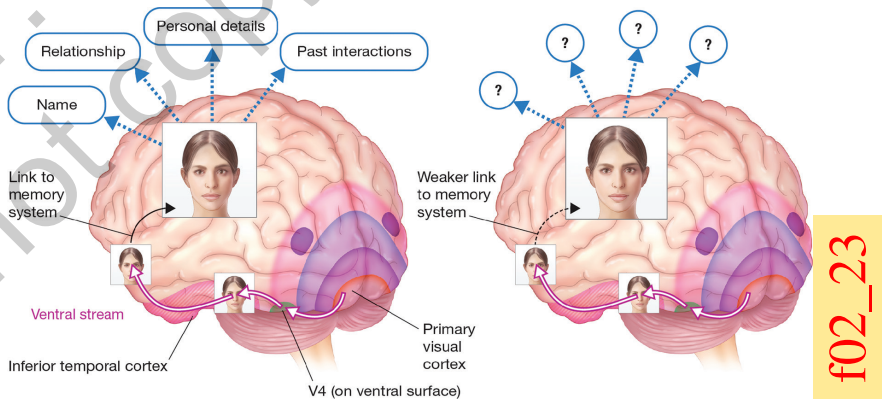
**FIGURE 2.22 ■ Training Eye Movements (A) and Reading Words (B).**



Source: "Neuro-visual rehabilitation" by N. Raz & N. Levin, 2017, *Journal of Neurology*, 264, 1051-1058. <https://link.springer.com/article/10.1007%2Fs00415-016-8291-0>

shown in the right panel of Figure 2.22A, requires them to track letters in an alphabetical sequence. Another training task, shown in Figure 2.22B, provides practice in reading words. The number of letters in the words increases while the presentation time decreases as training progresses. To perceive the entire word, patients are trained to fixate on either the beginning or the end of the word depending on their particular deficit. Training tasks also exist for large visual

**FIGURE 2.23 ■ Disruptive Pathways Causing Face Blindness.**



Source: Adapted from *INNATE: How the Wiring of Our Brains Shapes Who We Are*, by K. J. Mitchell, 2018, Princeton, NJ: Princeton University Press.

Copyright ©2023 by SAGE Publications, Inc.

fields. The person in the left panel of Figure 2.22A is searching for a square composed of four red dots.

In contrast to apperceptive agnosia, **associative agnosia** patients can combine visual elements into a whole perception but are unable to identify that perception. The most curious fact about these patients is they can accurately copy a line drawing but are unable to recognize what they have drawn! Essentially these patients can perceive the object but can no longer associate their perception with its meaning.

Face blindness provides an informative case study of how the “what” stream fails to connect to other parts of the brain. An area of the cortex known as the *fusiform face area* is responsive to recognizing that an object is a face, even for people with face blindness (Mitchell, 2018). Although the brain performs the initial stage of face processing perfectly well, it fails to communicate that information with the frontal cortex for people with face blindness (Figure 2.23). The link to the frontal cortex is necessary to recall information such as the person’s name, personal details, relationship, and past interactions.

The condition can be so debilitating that patients may not recognize close family members or even their own face. The famous neuropsychologist Oliver Sacks suffered from face blindness prior to his death. His book *The Man Who Mistook His Wife for a Hat* is based on a clinical case study of such a patient (Sacks, 1985). Interestingly, many patients suffering from this disorder can recognize the faces of loved ones after they hear them speak.

## SUMMARY

Pattern recognition is a skill that people perform very well. Three explanations of pattern recognition are template, feature, and structural theories. A template theory proposes that people compare two patterns by measuring their degree of overlap. A template theory has difficulty accounting for many aspects of pattern recognition. The most common theories of pattern recognition, therefore, assume that patterns are analyzed into features. Perceptual discrimination requires discovering distinctive features that distinguish between patterns. Treisman’s experiments on feature integration theory explored how a perceiver combines two features that are analyzed by separate parts of the visual system. Structural theories state explicitly how the features of a pattern are joined together. They provide a more complete description of a pattern and are particularly useful for describing patterns consisting of intersecting lines.

Sperling’s interest in the question of how many letters can be perceived during a brief exposure resulted in the construction of information-processing models for visual tasks. Sperling proposed that information is preserved very briefly in a visual information store, where all the letters can be simultaneously analyzed. When a letter is recognized, its name is verbally rehearsed and preserved in an auditory store that is a part of short-term memory.

Recognition of letters in a word is influenced by perceptual information and the letter context. The finding that a letter can be recognized more easily when it is part of a word than when it is part of a nonword or is presented by itself has been called the *word superiority effect*. An influential



model of this effect is the interactive activation model proposed by McClelland and Rumelhart. Its major assumption is that knowledge about the words of a language interacts with incoming feature information to provide evidence regarding which letters are in the word. Scenes are composed of interacting objects that are typically arranged in a meaningful spatial layout. Recognizing objects in scenes is often driven by accomplishing goals. Deep neural networks, used in scene recognition and many other complex AI tasks, utilize the same principles of simpler networks but have added multiple layers of connections to fine-tune the weights of thousands of connections.

*Visual agnosia* is a disruption in the ability to recognize objects. There are specialized forms of recognition disorders, such as an inability to recognize objects or familiar places. The “where” pathway is located in the upper parietal area and is primarily associated with object location and spatial attention. The “what” pathway supports object recognition and is primarily located in the lower temporal lobes. Patients with apperceptive agnosia are unable to combine visual features into a complete pattern whereas associative agnosia patients can, but these patients can not identify the pattern.

### RECOMMENDED READING

Hoffman's (1998) book, *Visual Intelligence*, provides both a readable and scholarly analysis of how we construct descriptions of objects. Fallshore and Schooler (1995) argue that verbally describing faces can lower later recognition because verbal descriptions ignore configural information. Kristjansson and Egeth (2019) provide an extensive history of how feature integration theory integrated relevant research in cognition, perception, and neuropsychology. Experts provide an overview of the theoretical contributions of neural networks (McClelland et al., 2010). For a history of neural networks that has resulted in the exciting accomplishments of deep networks read *The Deep Learning Revolution* (Sejnowski, 2018). A very readable

introduction to genetics and brain circuits is Kevin Mitchell's (2018) book *INNATE: How the Wiring of Our Brains Shapes Who We Are*.

DRAFT. NOT A FINAL PROOF.  
Do not copy, post, or distribute.



Getty Images/Yoshiyoshi Hirokawa



**ACTION**

**Action Joins Cognition**

Action Influences Perception  
 Actions Can Be Simulated  
 Gestures Influence Comprehension  
 Gestures Influence Thinking

**Combining Actions and Objects**

Physical Actions  
 Virtual Actions  
 Mental Actions

**Organizing Actions**

Creating High-Level From Low-Level Actions  
 Action Identification Theory  
 Neuroscience of Actions

**Applications**

Cognitive Offloading  
 Instruction

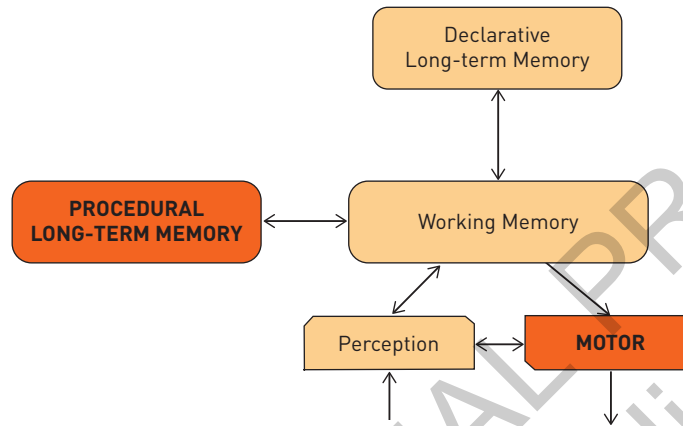
**SUMMARY****RECOMMENDED READING****LEARNING OBJECTIVES**

1. Describe how action influences perception and thinking.
2. Generate combinations of physical, virtual, and mental actions with physical, virtual, and mental objects.
3. Explain how focus can shift between low-level and high-level actions.
4. Identify common principles that apply to cognitive offloading and instruction.

The chapter on action discusses a different perspective from initial theories in cognitive psychology that emphasized the mental representation of knowledge. The initial orientation considered perception and action to be peripheral, not central, components of cognitive processing. The new perspective—**embodied cognition**—considers perception and action to be central components of cognition (Gibbs, 2006; Wilson, 2002). This chapter continues the discussion of procedural long-term memory (LTM) and adds the motor component of the standard model (Figure 6.1). Procedural LTM stores procedures for organizing actions, and the motor component executes the procedures.

**embodied cognition** A theoretical framework in which perception and action have a central role in cognition

**FIGURE 6.1** ■ The procedural LTM and motor components of the standard model of the mind.

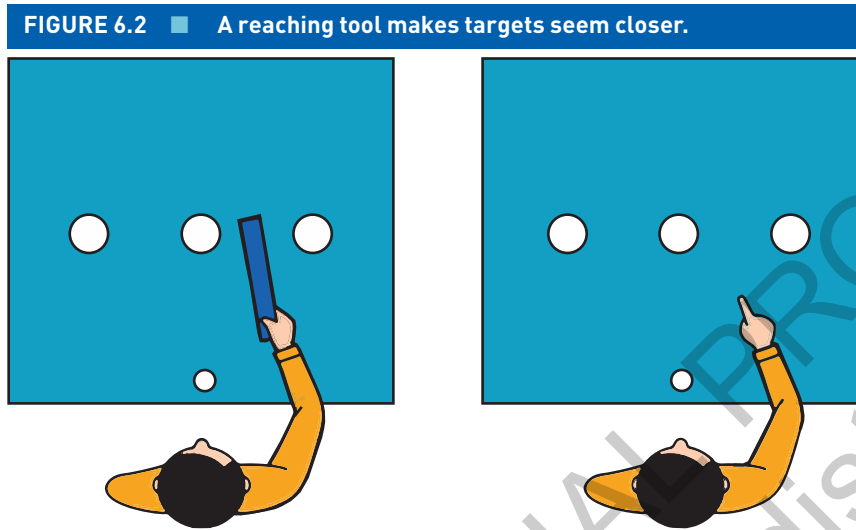


Source: Based on "A standard model of the mind: Toward a common computational framework across artificial intelligence, cognitive science, neuroscience, and robotics," by J. E. Laird, C. Lebiere, & P. S. Rosenbloom, 2017, *AI Magazine*, 38, 13–26. <https://doi.org/10.1609/aimag.v38i4.2744>

The embodied cognition perspective focuses on understanding cognition as an ongoing interaction with the external world. It emphasizes that much of thought is more than contemplation—it involves processes that determine possible actions. Although actions were also emphasized in the behaviorist movement described in Chapter 1, embodied cognition places action within a broader theoretical framework. This framework specifies that actions are (1) driven by goals that sometimes fail, (2) require planning and decisions among alternative actions, and (3) and involve prediction or anticipation of an intended outcome (Engel et al., 2013).

Goals, planning, and anticipation occur even for relatively simple actions, such as grasping an object (Rosenbaum et al., 2012). Consider a waiter filling water glasses that are initially placed upside down on the tables. To turn the glass upright, the waiter picks up the glass with his thumb down so he will be holding it with his thumb up after rotating it 180 degrees. Anticipation is also needed in social situations. It is polite to pass a water pitcher to another person so that person can grasp the pitcher by its handle.

The initial section of this chapter is about the interactions between actions and other components of cognition. It describes how actions influence perception, comprehension, and thinking. Actions typically act on objects, so the second section explains how physical, virtual, and mental actions act on physical, virtual, and mental objects. The third section examines how low-level actions, such as reaching for a water pitcher, become organized into high-level goals, such as filling a glass. The final section on applications discusses how actions can reduce cognitive load and support instruction.



Source: "Action's effect on perception," by J. K. Witt, 2011 *Current Directions in Psychological Science*, 20, 201–206. <https://doi.org/0.1177/0963721411410877>

## ACTION JOINS COGNITION

### Action Influences Perception

According to the standard model in Figure 6.1, cognitive components do not work in isolation but interact with each other. One consequence is that visual perception is not solely a visual process. What one sees is influenced not only by sensory information but by a person's purpose, physiological state, and emotions (Proffitt, 2006). For instance, studies have consistently found that the judged steepness of hills is influenced by physiological demands. Runners at the beginning of a race judge hills as less steep than at the end of a race, and more-fit people judge hills as less steep than those who are less fit. The advantage of this perceptual bias is that it simplifies planning. People who are tired or less fit require more energy to get up the hill.

This action-specific account of perception proposes that people perceive their surrounding environment in terms of their ability to act on it. Witt (2011) reviews extensive research documenting this effect, ranging from daily activities to superior athletic performance. Perceivers with narrow shoulders perceived doorways to be wider compared with perceivers with wide shoulders. Perceivers with a reaching tool estimated targets to be closer than perceivers who did not have the tool (Figure 6.2). Perceivers burdened with a heavy load judged distances as further and hills as steeper.

**amodal** Knowledge that is abstracted from sensory experiences

**modal** Knowledge is represented as sensory experiences



**PHOTO 6.1** Actions such as opening a drawer can be mentally simulated.

istockphoto.com/sirawit99

#### TABLE 6.1 ■ Object verification task

1. She pounded the nail into the floor.
2. She painted the fence with a brush.
3. She pounded the nail into the wall.



Source: "The effect of implied orientation derived from verbal context on picture recognition," by R. A. Stanfield & R. A. Zwaan, 2001, *Psychological Science*, 12, 153–156. <https://doi.org/10.1111/1467-9280.00326>

Previous successful performance makes a task appear easier. Softball players perceive the ball as larger when hitting well. Tennis players perceive the net as lower when returning balls successfully. Golfers perceive the hole as larger when playing better.

The action-specific perspective shares with Gibson's (1979) ecological approach its emphasis on the relation between perception and action. However, it differs from Gibson's theory of direct perception by demonstrating that perceivers interpret information differently in different circumstances. Witt and Riley (2014) proposed a reconciliation that makes the two perspectives more compatible. The reconciliation assumes that action-related information is not conceived of as an internal store of knowledge that relies on logical inferences but rather is detected at the time the action is anticipated. This perspective enables action-specific perception to incorporate



**PHOTO 6.2** Gestures accompany phone conversations.

istockphoto.com/fizkes

concepts such as intention, attention, and information selection that are considered in ecological psychology.

### Actions Can Be Simulated

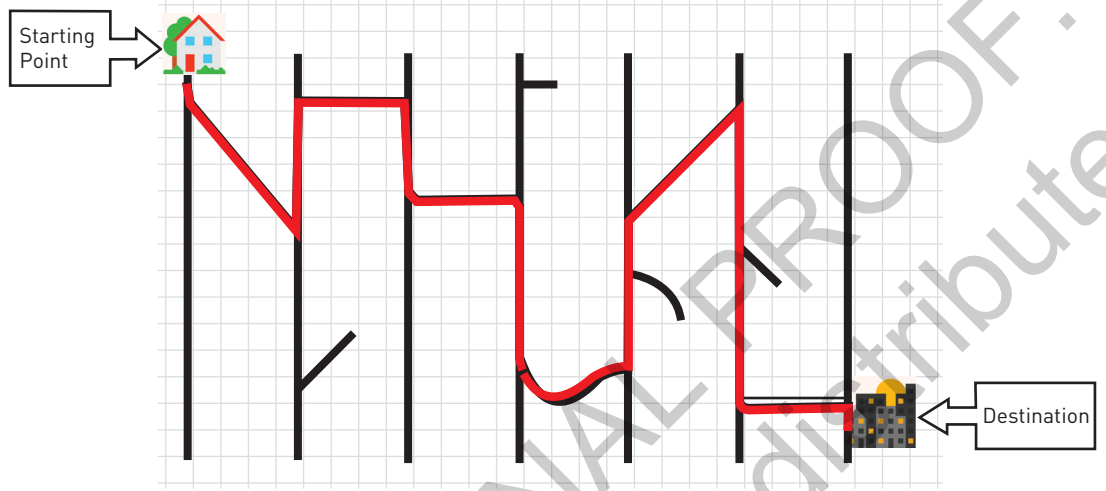
An extensive article by Larry Barsalou (1999) at Emory University helped establish the embodied cognition framework. Barsalou argued that theories at that time were primarily **amodal** because they did not directly represent the perceptual experiences encountered in learning about concepts. A different theoretical approach is the perceptual symbols system proposed by Barsalou (1999), in which perceptual experiences are directly stored in LTM. A perceptual symbols system provides a **modal** framework that directly stores sensory experiences such as audition, vision, taste, smell, and touch. These sensory experiences can be re-experienced by mentally simulating them. For instance, try to imagine the smell, taste, and texture of a warm apple pie that was recently removed from the oven. People describe attractive foods by emphasizing words referring to taste and texture. They describe neutral foods by using words referring to visual appearance (Papies et al., 2020).

There is evidence that people mentally simulate actions. In one experiment, college students at the University of Wisconsin had to quickly judge whether a phrase, such as “open the drawer” or “boil the air,” made sense (Glenberg & Kaschak, 2002). They had to move their hand either

---

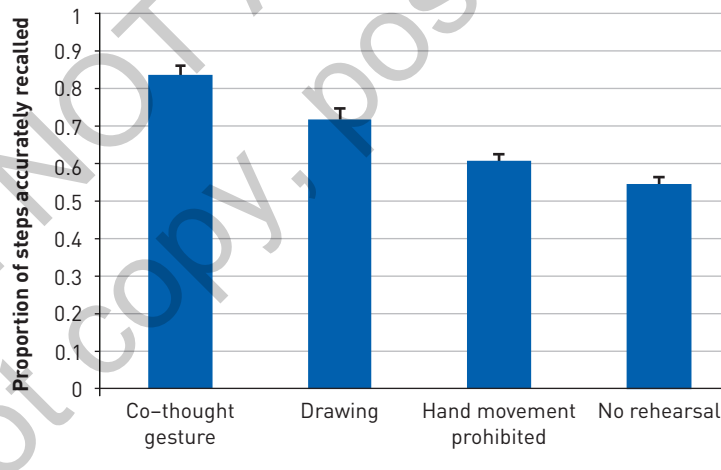
**penetrative thinking** The ability to reason about the interior structure of an object based on what is visible on its surface



**FIGURE 6.3** ■ The red line shows a sequence of steps during route navigation.

*istockphoto.com.photosynthesis*

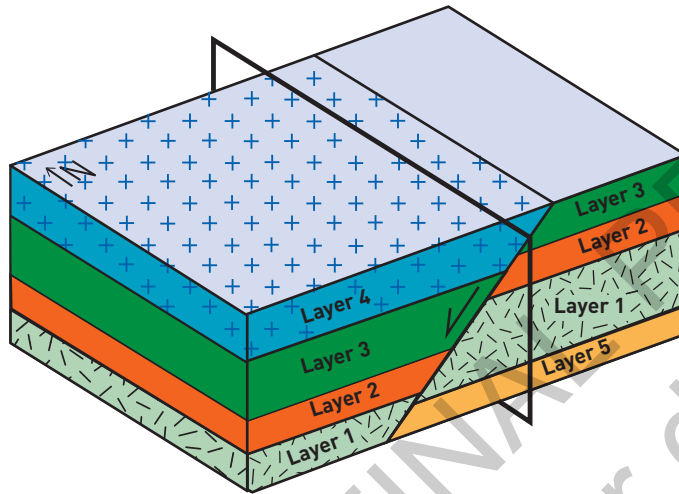
Source: "Producing gestures facilitates route learning," by W. C. So, T. H.-W. Ching, P. E. Lim, X. Cheng, & K. Y. Ip, 2014, *Plos One*, 9(e112543). <https://doi.org/10.1371/journal.pone.0112543>

**FIGURE 6.4** ■ Mean proportion of correctly recalled steps across four learning conditions.

Source: "Producing gestures facilitates route learning," by W. C. So, T. H.-W. Ching, P. E. Lim, X. Cheng, & K. Y. Ip, 2014, *Plos One*, 9(e112543). <https://doi.org/10.1371/journal.pone.0112543>

toward or away from their body to hit a response key. Participants were faster in responding when the response required an action in the same direction as the one implied in the

**FIGURE 6.5** ■ Image of a block diagram similar to diagrams shown to participants in both the gesture and gesture-prohibited groups.



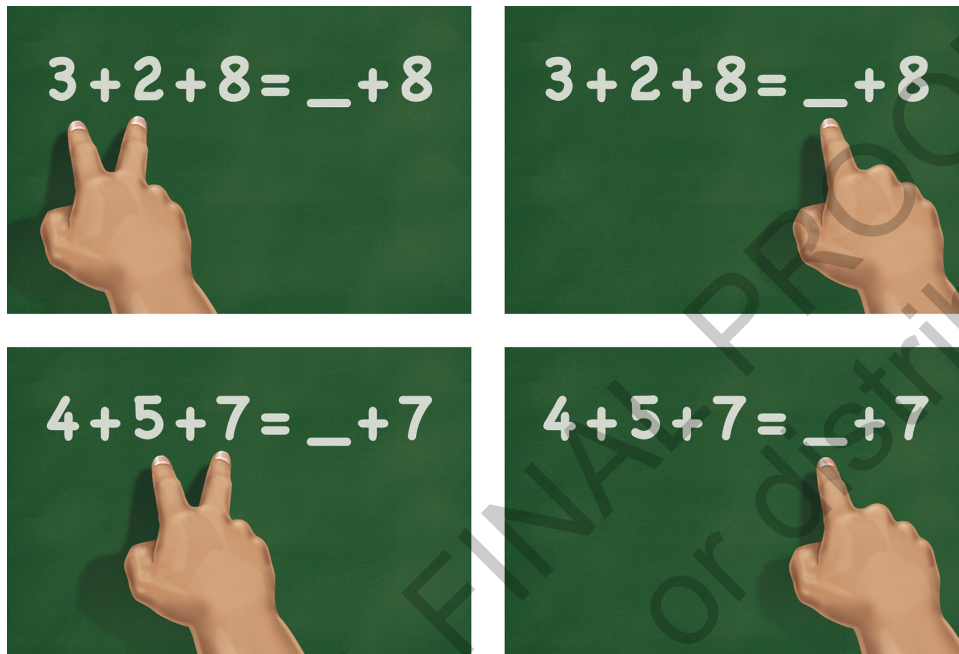
Source: "Student gestures aid penetrative thinking," by K. Atit, K. Gagnier, & T. F. Shipley, 2015, *Journal of Geoscience Education*, 63, 66–72. <https://doi.org/10.5408/14-008.1>

phrase. For example, they were faster in responding by moving their hand toward their body when verifying the statement "open the drawer" and faster when moving their hand away from their body for the statement "close the drawer." These findings are consistent with a mental simulation of the action that requires moving your arm toward your body to open a drawer and away from your body to close a drawer (Photo 6.1).

Results obtained by Stanfield and Zwaan (2001) supported the hypothesis that simulations help people comprehend the meaning of verbally described actions. For example, reading that someone put a pencil in a drawer should evoke an image of a horizontal pencil, and reading that someone put a pencil in a cup should evoke an image of a vertical pencil. Stanfield and Zwaan tested their hypothesis that visual simulations of verbal statements would include an object's orientation by asking students at Florida State University to quickly decide whether a pictured object had been mentioned in a sentence that they just read. You can obtain an approximate idea of this task by responding to whether the object in Table 6.1 is mentioned in each of the three sentences. This demonstration is approximate because the test object did not appear until after participants read each sentence in Stanfield and Zwaan's experiment.

According to the visual simulation hypothesis, the time to verify a mentioned object should depend on whether the picture matches the implied orientation in the sentence. The results supported the hypothesis, as illustrated by the distinction between the two sentences: *She pounded the nail into the floor*, versus *she pounded the nail into the wall*. The readers were faster in confirming a picture of a vertical nail following the first sentence (pounding a nail into the floor)

**FIGURE 6.6** ■ Correct gesture (above) and partially correct gesture (below).



Source: "Action to abstraction: Gesture as a mechanism of change," by S. Goldin-Meadow, 2015, *Developmental Review*, 38, 167–184. <https://doi.org/10.1016/j.dr.2015.07.007>

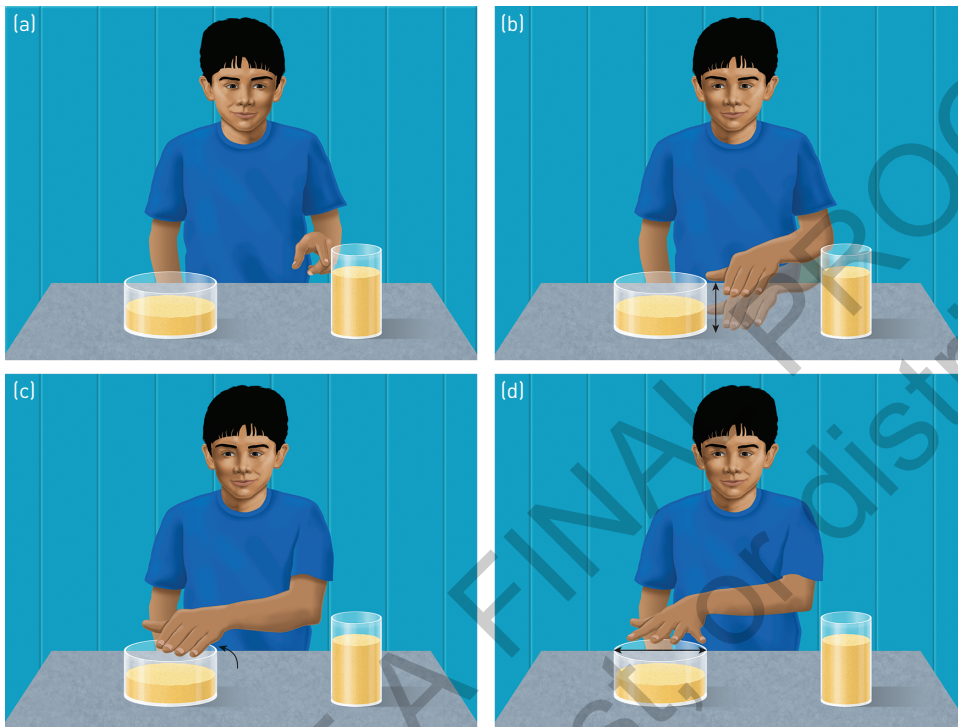
and were faster in confirming a picture of a horizontal nail following the third sentence (pounding the nail into the wall).

### Gestures Influence Comprehension

One advantage of mentally simulating actions during comprehension is that we can use these simulations as a basis for gestures. In other words, the same simulations can provide the foundation for both comprehending and producing language. This is exactly the claim made by Hostetter and Alibali (2008) in their gestures-as-simulated action framework.

As proposed in embodied theories of cognition, the link between perception and action is central. The simulations that occur during language comprehension are based on both visual and motor images. Motor imagery activates premotor areas in the brain that have the potential to spread to motor areas and create gestures. The gestures-as-simulated action framework proposes that three factors determine whether a gesture occurs. The first is the strength of the simulated action. Some simulations involve only visual imagery. The word "beautiful" will likely invoke visual imagery but not action imagery. The second factor is the height of the gesture threshold. You probably know some people who have a low threshold and use their hands frequently as they speak and others who have a high threshold and rarely gesture. The third factor

**FIGURE 6.7** ■ Using gestures for exploration during problem-solving.



Source: "How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis," by S. Kita, M. W. Alibali, & M. Chu, 2017, *Psychological Review*, 124, 245–266. <https://doi.org/10.1037/rev0000059>

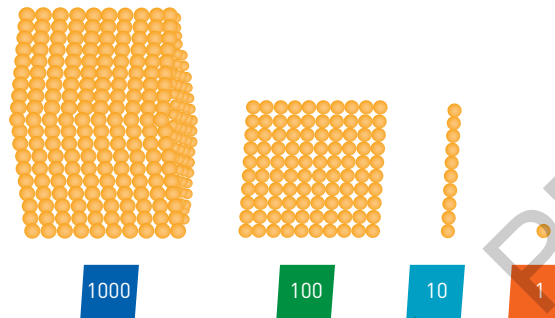
**TABLE 6.2** ■ Combinations of physical, virtual, and mental actions and objects

Actions	Objects		
	Physical	Virtual	Mental
Physical	Montessori Manipulatives	Wii Sports Games	Gestures
Virtual	Robotic Surgery	Virtual Experiments	Teaching the Blind
Mental	Brain-computer Robotic Interfaces	Brain-Computer Cursor Interfaces	Sports Simulations

Source: "Combining physical, virtual, and mental actions and objects," by S. K. Reed, 2018, *Educational Psychology Review*, 30, 1091–1113. <https://doi.org/10.1007/s10648-018-9441-y>

is speech. Although both comprehending and producing language depend on simulations, gestures typically occur only when people produce language.

**FIGURE 6.8** ■ Numerical tiles help children connect an abstract representation to the more concrete beads.



Source: "What makes mathematics manipulatives effective? Lessons from cognitive science and Montessori education," by E. V. Laski, J. R. Jordan, C. Daoust, & A. K. Murray, 2015, *SAGE Open*, April–June, 1–8. <https://doi.org/10.1177/2158244015589588>

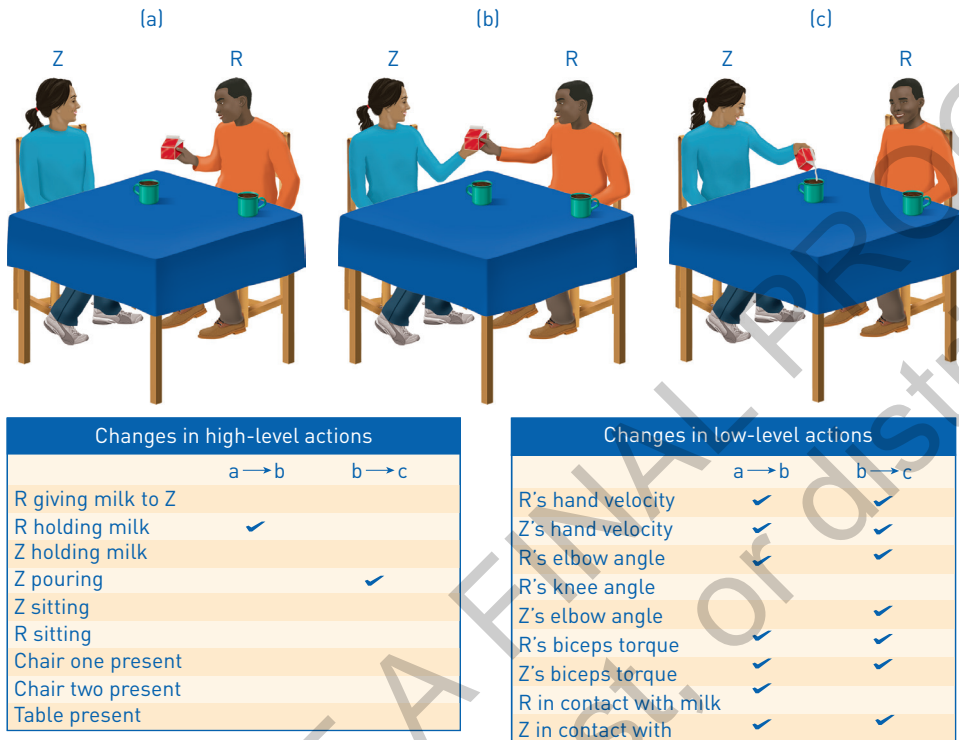


**PHOTO 6.3** Robotic equipment to perform minimally invasive surgery.

istockphoto.com/MARHARYTA MARKO

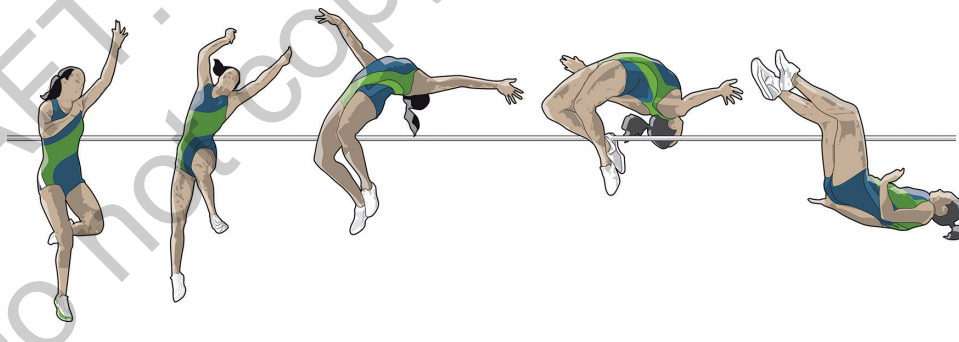
The number of gestures that speakers produce depends on the situation. For instance, if we use gestures to help people understand, then we should gesture more when talking to them face-to-face than when talking to them over the telephone. A group of researchers tested this

**FIGURE 6.9** ■ Changes in high-level (left) and low-level (right) actions.



Source: "Constructing experience: Event models from perception to action," by L. L. Richmond & J. M. Zacks, 2017, *Trends in Cognitive Sciences*, 21, 962–980. <https://doi.org/10.1016/j.tics.2017.08.005>

**FIGURE 6.10** ■ The Fosbury Flop.



Dorling Kindersley / Getty Images



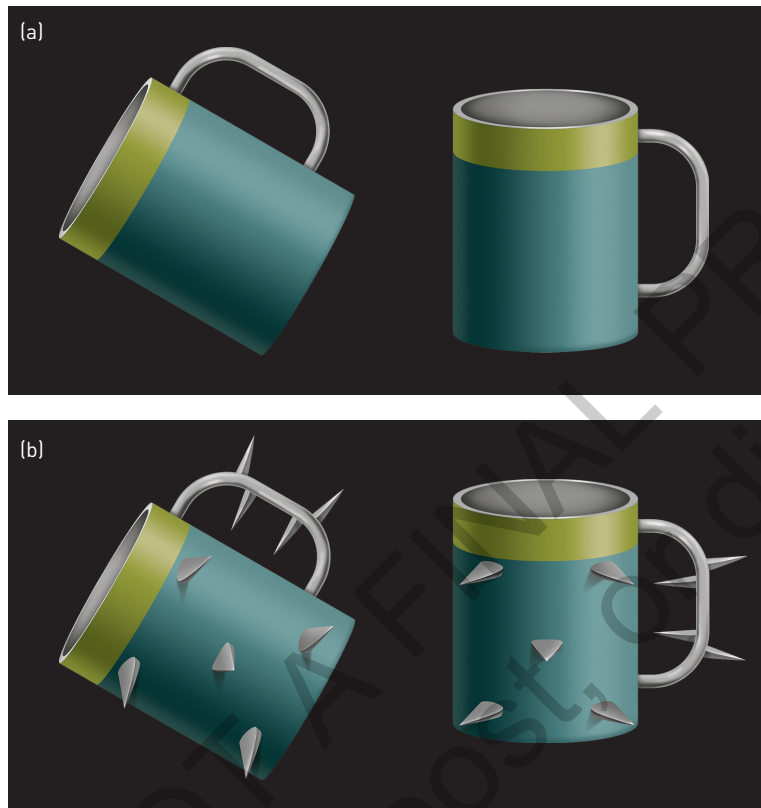
**PHOTO 6.4** Focusing on basic actions improves performance.

JJ pixs / Alamy Stock Photo

hypothesis in a study that varied whether the listener could observe the gestures (Bavelas et al., 2008). Some participants described two pictures in a face-to-face dialogue to another person who could not see the pictures. Other participants described the pictures to another person in a telephone conversation. A third group described the pictures in a monologue to a tape recorder. The average number of gestures declined from an average of 21.8 per minute in the face-to-face conversation to 14.9 per minute in the telephone conversation to 4.3 per minute in the monologue. Speakers in the face-to-face condition were also more likely to put information in their gestures that was not in their words and refer to their gestures during the conversation. But perhaps the most surprising aspect of these findings is the large number of gestures made in a telephone conversation, suggesting that gestures aid speakers even when they know the listener cannot observe them (Photo 6.2).

This finding is consistent, nonetheless, with studies that reveal the effect of producing gestures is greater than the effect of observing gestures for comprehending speech (Dargue et al., 2019). The majority of studies used recall as a measure of comprehension, which was defined as an individual's understanding of a verbal message, such as a narrative or a set of verbal instructions. A possible explanation of the beneficial effect of producing gestures is that the speaker's gestures reduce cognitive load, enabling more resources for understanding. We will see evidence for this hypothesis later in the chapter. Other findings revealed that gestures were equally effective across a wide variety of age groups and measures of recall (Dargue et al., 2019).

FIGURE 6.11 ■ Smooth versus spiky cups.



Source: "Co-thought and co-speech gestures are generated by the same action generation process," by M. Chu & S. Kita, 2016, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42, 257–270. <https://doi.org/10.1037/xlm0000168>

### Gestures Influence Thinking

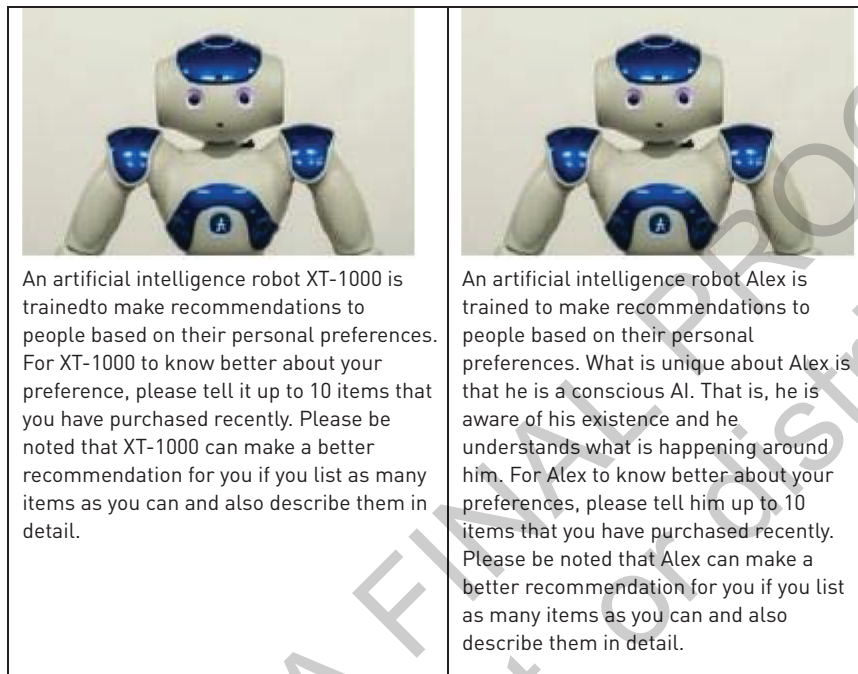
The gesture-for-conceptualization-hypothesis (Kita et al., 2017) proposes that gestures activate, manipulate, package, and explore spatio-motoric information for both speaking and thinking. This section provides examples of research to illustrate how each of the following principles applies to thinking:

1. Gestures activate and maintain spatio-motoric information.
2. Gestures manipulate spatio-motoric information.
3. Gestures package spatio-motoric information.
4. Gestures explore spatio-motoric information.

Evidence that gestures facilitate the learning of routes shows how they *activate and maintain* spatio-motoric information (So et al., 2014). Undergraduates viewed diagrams in which a red line displayed a sequence of steps to reach a destination (Figure 6.3). In the co-thought gesture

Copyright ©2023 by SAGE Publications, Inc.



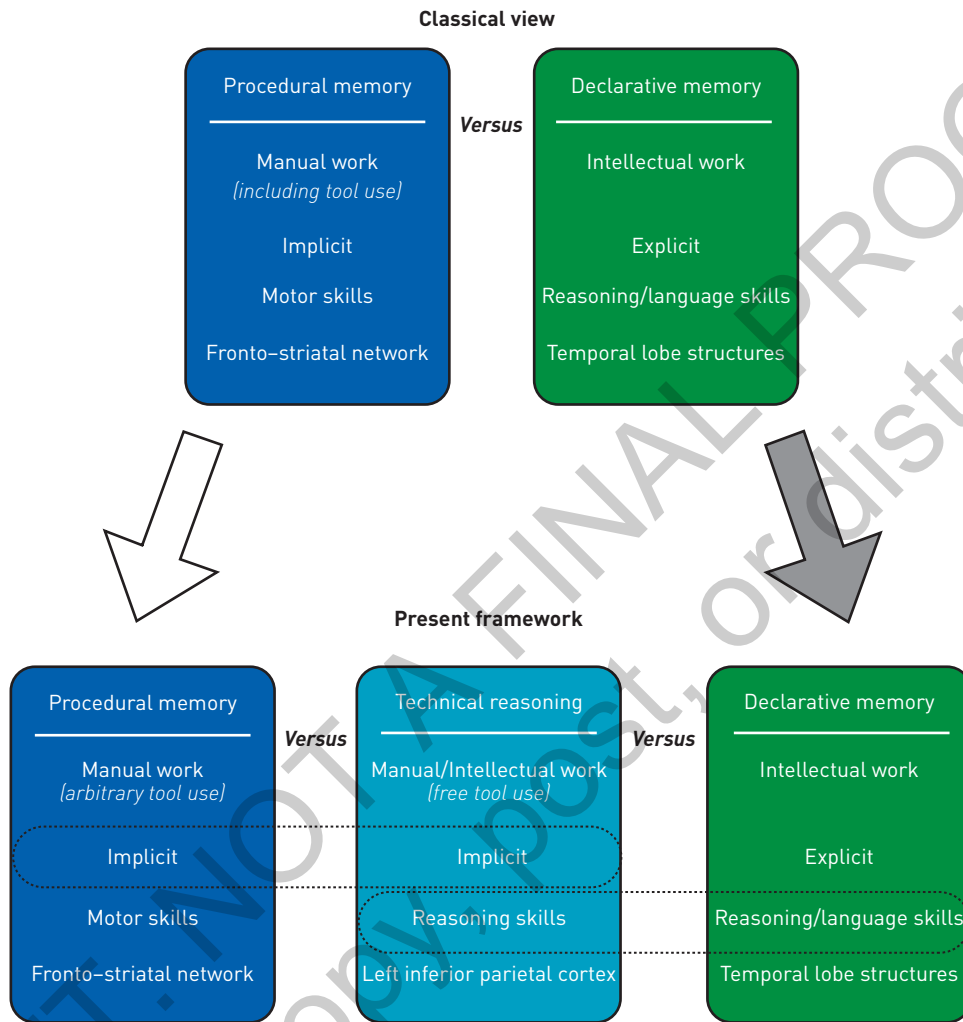
**FIGURE 6.12** ■ A non self-aware (left) versus a self-aware robot (right).

Source: Kim, T. W., & Duhachek, A. (2020). Artificial intelligence and persuasion: A construal-level account. *Psychological Science*, 31, 363–380. <https://doi.org/10.1177/0956797620904985>

**FIGURE 6.13** ■ A low-construal (left) versus a high-construal message (right).

Source: From Kim, T. W., & Duhachek, A. (2020). Artificial intelligence and persuasion: A construal-level account. *Psychological Science*, 31, 363–380. <https://doi.org/10.1177/0956797620904985>

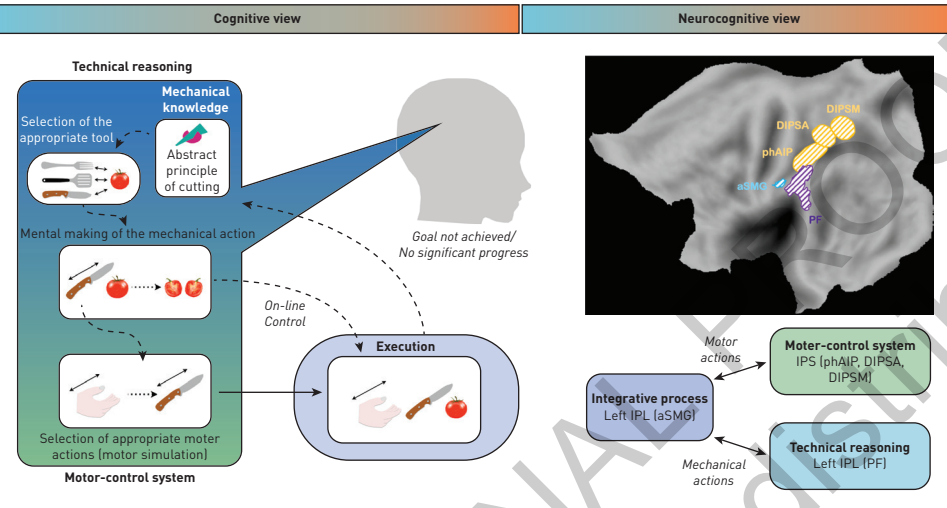
condition, participants were instructed to rehearse the routes while using their hands. In the drawing condition, they drew the route. In the hand-movement-prevention condition, they

**FIGURE 6.14** ■ A unified framework for studying human tool use.

Source: "Looking for intoelligence: A unified framework for the cognitive study of human tool use and technology," by F. Osiurak & D. Heinke, 2018, *American Psychologist*, 73, 169-185. <https://doi.org/10.1037/amp0000162>

visualized the route without moving their hands. In the no-rehearsal (control) condition, they performed a distracting task. As shown in Figure 6.4, the proportion of accurately recalled steps was significantly higher when using gestures during learning, even higher than drawing the route.

Gestures also enable us to demonstrate how we would *manipulate* material, as studied in an evaluation of **penetrative thinking**, defined as the ability to reason about the interior structure of an object based on what is visible on its surface. Penetrative thinking is needed to be

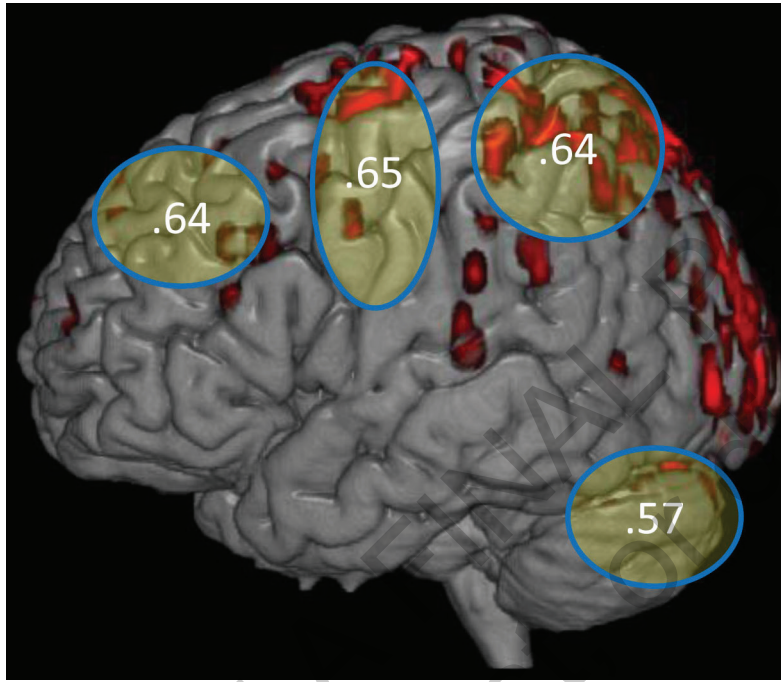
**FIGURE 6.15** ■ A novel task requires both declarative and procedural knowledge.

Source: "Technition: When tools come out of the closet," by F. Osiurak, M. Lesourd, J. Navarro, & E. Reynaud, 2020, *Perspectives on Psychological Science*, 15, 880–897. <https://doi.org/10.1177/1745691620902145>

**FIGURE 6.16** ■ A nautical knot used to study the neural representations of imagined actions.

Source: Mason, R. A., & Just, M. A. (2020). Neural representations of procedural knowledge. *Psychological Science*, 31, 529–540.

**FIGURE 6.17** ■ Brain regions that distinguished among different imagined actions.



Source: From Mason, R. A & Just, M. A. (2020). Neural representations of procedural knowledge. *Psychological Science*, 31, 529-540.

successful in many scientific disciplines, particularly the geosciences. Figure 6.5 shows a block diagram used to support geoscience instruction (Atit et al., 2015). Participants in a gesture group were told “using your hands, can you show me how you would build the structure from flat layers of Play-Doh”? They were next asked to imagine slicing the block at the bold black line in the diagram and explain what layers they would see in the cross-section. Participants in a gesture-prohibited group received the same instructions but could not use their hands during their explanations. Only the gesture group improved their scores on the Geologic Block Cross-Sectioning Test following the completion of three exercises. The researchers recommended that gesture exercises be included during instruction to help reason and communicate about spatial information.

The *packaging* of information refers to creating units of information that are useful for performing a current task (Goldin-Meadow et al., 2009). Figure 6.6 shows two methods for packaging information in an addition task. In the upper two screens, a student uses a correct  $V$  gesture to group two numbers ( $3 + 2$ ) and a pointing gesture to demonstrate that their sum fits into the blank on the right side of the equation. In the lower two screens, a student uses a partially correct  $V$  gesture to group two numbers ( $5 + 7$ ). Their sum does not fit into the blank, but the gestures



**PHOTO 6.5** Using notes to offload information onto the environment.

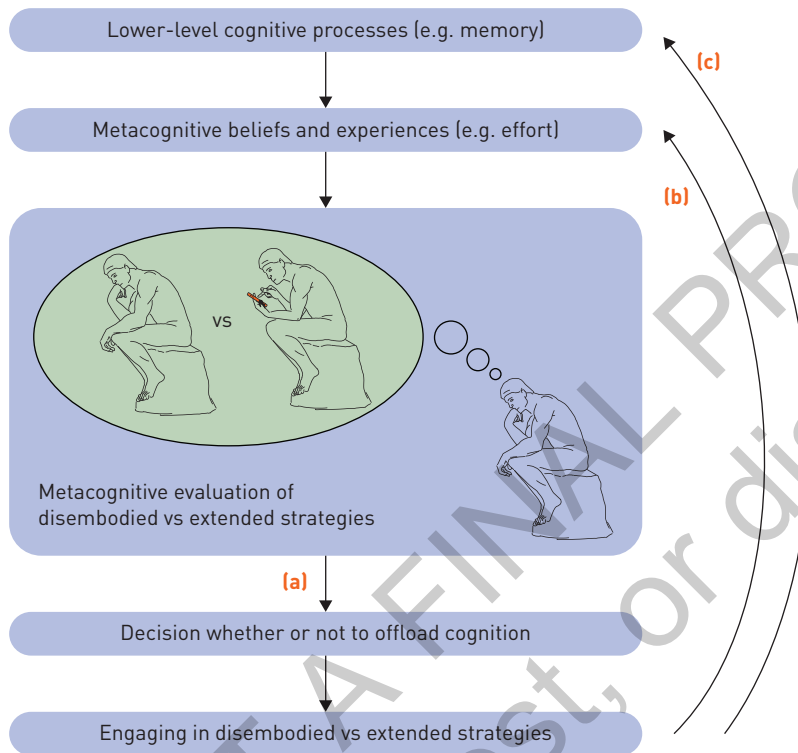
istockphoto.com/diego\_cervo

illustrate that numbers can be grouped and inserted into the right side of the equation. The students required to produce correct gestures learned more than children required to produce partially correct gestures, who learned more than children required to produce no gestures.

Problem-solving initiates the last principle—the use of gestures in *exploration*. Consider a Piagetian task in which the experimenter pours the contents of one of two identical tall glasses of sand into a short, wide dish (Alibali et al., 2000). The boy in Figure 6.7 initially stated that the remaining tall glass contained more sand than the short dish. While explaining his judgment, he initially focused on the tall glass (a) but decided against it. He then explored the width and height of the dish with an up and down gesture at the side of the dish (b). He next moved his hand over the dish (c) and explored the area by spreading his hand (d). This exploration helped him generate the idea that the sand fills a larger area in the dish that could compensate for its lack of height.

A difference between gestures and action is that gestures represent the world but do not physically change it. Gestures promote understanding because they bring action into mental representations at a higher level of abstraction than manipulation. As stated by Goldin-Meadow (2015, p. 170):

Gestures are representational and thus more abstract than direct actions on objects. It may be this comfortable middle ground, with one foot in concrete action and one foot in abstract representation, that makes gesture such a powerful tool for learning.

**FIGURE 6.18** ■ Decision regarding cognitive off-loading.

Source: "Cognitive offloading," by E. F. Risko & S. J. Gilbert, 2016, *Trends in Cognitive Sciences*, 20, 676–688. <https://doi.org/10.1016/j.tics.2016.07.002>

## COMBINING ACTIONS AND OBJECTS

Gestures occur at a higher level of abstraction than manipulation because they do not manipulate physical objects. We can use our hands to demonstrate how we would physically pound a nail into the floor, but the hammer and nail are imaginary or mental objects. Previously, we saw evidence for imagining this situation in which both the action and object are mental. These situations are examples within a taxonomy that combines physical, virtual, and mental actions with physical, virtual, and mental objects (Table 6.2). Let's explore the taxonomy in the next sections on physical, virtual, and mental actions.

### Physical Actions

Many readers have some knowledge about the emphasis on physical manipulatives in Montessori schools. In her book *Montessori: The Science Behind the Genius*,

Angeline Lillard (2005) discusses Montessori's work and subsequent research that has supported many of Montessori's ideas about learning and development. A review of research on

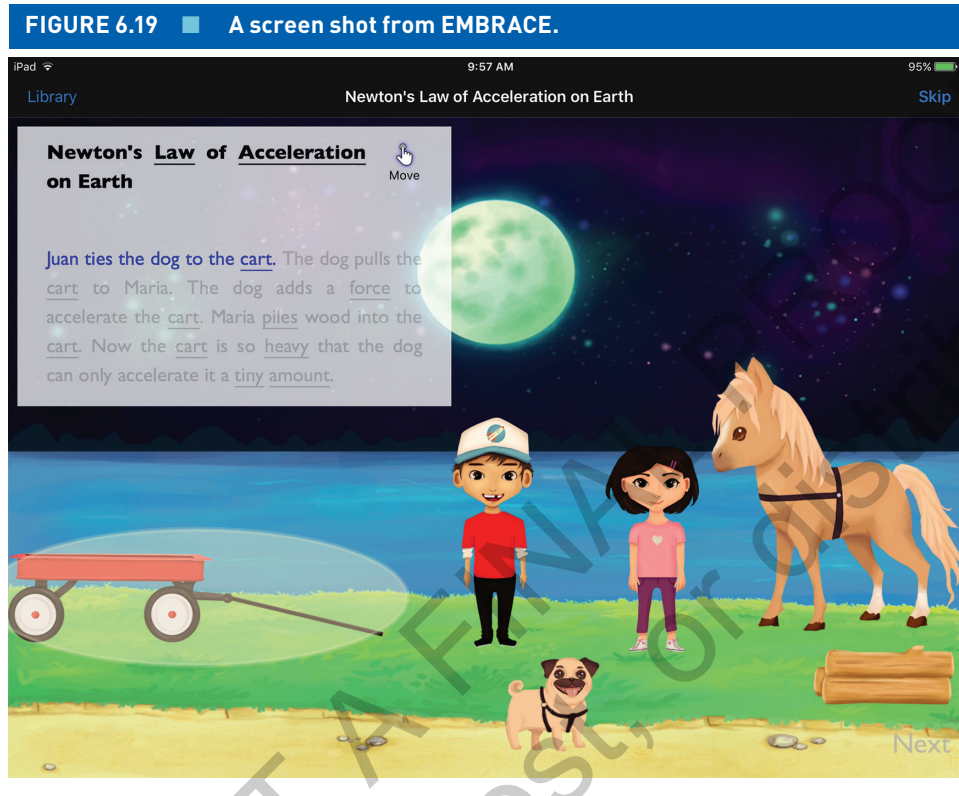


Figure courtesy of Arthur Glenberg.

Source: A screen shot from EMBRACE, an iPad learning app.

what makes mathematics manipulatives effective contained four recommendations (Laski et al., 2015). The first is to use manipulatives consistently over a long period of time. For instance, the Montessori golden bead materials are used throughout the early elementary years to help children develop an understanding of the base-10 system. The manipulatives represent the base-10 number system by individual beads that can be assembled into 10 connected beads that form a 10 by 10 square of 100 square beads that comprise a cube of 1000 beads. The second recommendation is to begin with concrete representations and move to more abstract representations. Figure 6.8 shows the replacement of the concrete beads with more abstract numerical tiles that can be used without the physical representation of the quantities. The third recommendation is to avoid manipulatives that have distracting irrelevant features. Because the beads are all the same color and size, children are not distracted by irrelevant attributes. The fourth recommendation is to explicitly explain the relation between a manipulative and a mathematical concept to help children make this connection.

The next combination in Table 6.2 pairs physical actions with virtual objects. Physical actions on virtual objects became very popular when Nintendo launched the video game console

Wii in 2006. The tracking of the Wii remote in three-dimensional space enabled players to use physical actions to control a virtual game on a screen (Sparks et al., 2009). Games included baseball, bowling, boxing, golf, and tennis. Although an innovative design, the Wii console had a higher rate of associated injuries than traditional game consoles.

Gestures offer a third combination—a physical action on a mental object. Although gestures contain many components of the actions they mimic, they also eliminate components. The force needed to lift an object is missing in a gesture that lifts nothing. However, gestures offer a potential advantage for transfer to other objects because an action such as lifting is not linked to a particular object (Goldin-Meadow, 2015).

### Virtual Actions

I classify robotic surgery as performing virtual actions on physical objects because the surgeon sits at a console (enabling virtual actions) to operate on the body (a physical object). The robotic da Vinci System was designed to improve upon conventional laparoscopy surgery in which the surgeon operates while standing, using hand-held, long-shafted instruments (Photo 6.3). The da Vinci System consists of a console that the surgeon uses to control four interactive robotic arms. Three of the arms hold instruments such as scalpels and scissors. The instruments exceed the natural range of motion of the human hand while motion scaling and tremor reduction further refine the surgeon's hand movements. Performing surgical movements on the computer console requires training unique actions that differ from those used in either open or laparoscopic surgery.

Surgery requires performing very skilled actions, but training in a virtual environment can be equivalent to training in a physical environment for learning concepts that do not require skilled actions. Lara Triona and David Klahr at Carnegie Mellon University-trained fourth- and fifth-grade students to design experiments to measure how far springs stretched for various combinations of springs and weights (Triona & Klahr, 2003). The springs could be long or short, wide or narrow, and thin or thick. The results showed that children who trained with virtual materials by varying only one variable at a time were as capable of designing good experiments as children who trained with physical materials. The virtual group also did as well as the physical group on a transfer task to evaluate the effect of steepness, length, surface, and type of ball on the time it would take a physical ball to roll down a physical ramp. The authors proposed that there are many advantages of computer-based laboratories, including portability, safety, cost-efficiency, and flexible, dynamic data displays (Klahr et al., 2007).

Virtual actions on mental objects occur when someone uses a virtual environment to create mental representations. A video game to teach navigation skills to the blind satisfies this requirement. Unlike sighted individuals, the blind must rely on nonvisual information to navigate the environment. An Audio-based Environment Simulator of a virtual building prepared the blind participants to subsequently navigate an actual building (Merabet et al., 2012). Exploration of the virtual building occurred through simple keystrokes while audio information described the location within the building. After training, the learners were evaluated on navigating a series of predetermined paths in the targeted physical building. They were highly successful, including



finding shortcuts within the building and the shortest path to exit the building from different starting points (Connors et al., 2014).

### Mental Actions

Mental actions on physical objects occur whenever someone uses thoughts to manipulate objects in the environment. The use of brain potentials to control robotic actions on physical objects requires the intervention of a robot to carry out the action between the thought and the object. Invasive sensors use surgical implants of electrodes; noninvasive sensors record brain signals from the scalp. Both approaches are based on the principle of cortical preparation that occurs before a cognitive, motor, or emotional response. Cortical preparation can be measured as a voltage shift in EEG (electroencephalographic) activity that can be used to control a physical device.

A team at the University of Palermo (Italy) used EEG recordings from the scalp to benefit amyotrophic lateral sclerosis (ALS) patients. Four ALS patients and four healthy controls learned to use the technology to control a robot to reach and grasp a glass of water. The brain computer interface (BCI) consisted of two high-level commands (grasp and give), four directional commands (forward, backward, left, right), and two turn commands. A few minutes of training was sufficient for enabling all four healthy participants and three of the four ALS patients to control the robot's actions at a high level of accuracy (Spataro et al., 2017).

The mental actions that control robots can also manipulate objects on a computer screen. However, moving a cursor occurs along two dimensions, and a single-modality EEG signal can only exert control along a single dimension. A group at the University of Electronic Science and Technology of China found that imagining more than one signal creates more natural two-dimensional diagonal movements (Ma et al., 2017). Horizontal movement occurs by imagining the movement of either the left or right hand. Vertical movement occurs by imagining the number 1 or 2. Imagining the four combinations LEFT 1, LEFT 2, RIGHT 1, and RIGHT 2 results in greater efficiency by controlling diagonal movements of the cursor.

Mental actions on mental objects occur during visual simulations that can be used for training in many domains. In their extensive review of the role of imagery in sports Cumming and Williams (2014) discuss the variables that influence the effects of imagery training on performance. Two commonly discussed attributes are vividness and controllability. Vividness refers to the clarity and sensory richness of the image. Controllability refers to the transformation and maintenance of a generated image. For instance, the ability to mentally change the viewing angle is helpful in domains such as sports and dance.

## ORGANIZING ACTIONS

---

### Creating High-level From Low-level Actions

Jeffrey Zacks has been a leader in constructing event models of how people perceive, remember, think about, and respond to events. An article by Richmond and Zacks (2017) provides a thorough summary of work on this topic. Some of this research investigates how low-level

actions become organized into high-level actions. As indicated by the checks in Figure 6.9, high-level (event model) actions begin with R holding the milk and end with Z pouring the milk. Low-level actions enable high-level actions and are described in terms of hand velocity, joints, muscle torques, and contact relations. These low-level actions support high-level actions, such as holding, getting, and pouring milk. High-level actions result in smoother changes in behavior and are more learnable. They also allow us to predict behavior by extrapolation from previous actions.

Most movement tasks involve a sequence of low-level actions that require integration to carry out complex skills (Diedrichsen & Kornysheva, 2015). A tennis serve consists of throwing the ball, taking a backswing, and accelerating the arm forward. Each of these phases involves the coordination of multiple body parts. Two stages in this coordination are the selection of particular goals and their execution through activities of the muscles. Although execution enables the performance of a complex skill, it does not support transfer to similar skills. Diedrichsen and Kornysheva (2015) therefore proposed an intermediate level between the execution and selection levels. The intermediate level enables a pianist to execute a specific chord transition within a new context and supports transfer between hands.

Gray and Lindstedt (2017) proposed that learning new methods during skill acquisition can improve performance but initially may cause a decline. An example is a shift from typing while looking at the keys to typing without looking at the keys. The average speed of experienced visually guided typists is approximately 30–40 words per minute. The shift to touch-typing initially results in a decline in performance, but following a long period of training, touch typists reach an average speed of 60 to 70 words per minute (Yechiam et al., 2003). The initial dips in performance are caused by learning new methods that are either mastered or abandoned by returning to methods that have worked in the past (Gray & Lindstedt, 2017).

The development of an innovative high-jumping technique by Dick Fosbury illustrates the power of innovation in improving motor skills (Gray & Lindstedt, 2017). The traditional technique required the jumper to attempt to clear the bar with his head facing the bar while lifting one leg at a time over the bar. The “Fosbury Flop” requires the jumper to go over the bar backward while kicking both legs in the air at the end of the jump (Figure 6.10). Fosbury demonstrated the effectiveness of this new technique when he won the gold medal and set an American record in the 1968 summer Olympics. The Fosbury Flop is now the most widely used method in high jumping.

### Action Identification Theory

In 1987, Robin Vallacher and Daniel Wegner published an influential article that has proven very helpful in providing a framework for understanding the organization of actions. Their action identification theory (Vallacher & Wegner, 1987) represents an action at different levels of specificity that range from low-level movements to a high-level goal. For instance, the act of calling to schedule an appointment can be identified as moving a finger, touching numbers, entering a phone number, talking to a person, and scheduling the appointment. Action identification theory contains four principles to capture the interdependence between these different levels of detail (Vallacher, 2007; Vallacher & Wegner, 1987).

1. Action is maintained with respect to its consciously available identity.
2. When both a lower- and a higher-level identity are available, there is a tendency for the higher-level identity to become conscious.
3. When an action cannot be maintained at its conscious identity, there is a tendency for a lower-level identity to become conscious.
4. The principles of the theory work together to promote the level that is most appropriate or optimal for performing the action.

The first principle—maintaining an action at a specific level—is needed to account for the stability of actions. Context, however, can influence the perceived level. Solving a mathematics puzzle could be perceived as “keeping track of numbers” or “performing mental calculations” in the privacy of one’s home but as “demonstrating my math skill” or “trying not to embarrass myself” in a testing situation.

The second principle—focusing on a higher level—provides meaning and implications for performing the action. Identities at a higher level have greater potential for creating one’s self-concept than identities at a lower level. Creating a piece of art provides more information about the person, for example, than does a low-level identity such as moving a paintbrush.

The third principle states that switching to a lower level occurs when it is difficult to maintain an action at a higher level. A novice tennis player may devise a strategy to win the match but soon discover that a focus on basic actions, such as preparing the racket and following through with the stroke, is necessary to improve performance (Photo 6.4). For actions as diverse as video games and playing the piano, people with greater experience were more likely to perceive their actions at the more meaningful higher levels.

Making the task more difficult also produces shifts in levels. An example is a study in which participants drank either from a normal coffee cup or one that was abnormally large. Participants who drank from a normal cup endorsed high-level actions, such as “getting energized” and “promoting my caffeine habit.” Participants who drank from the unwieldy cup endorsed low-level actions, such as “lifting a cup to my lips” and “swallowing.”

The last principle states that the first three principles work together to determine the most appropriate level of performance (Vallacher, 2007). Although people focused on low-level actions to drink from an abnormally large cup, they avoided thinking about low-level actions when they imagined rotating a spiky cup (Figure 6.11). Participants were shown different pairs of mugs and asked to describe how the left mug in each pair could be rotated to the position of the right mug. The average number of words used to describe the rotation did not differ between the smooth and spiky mugs. However, participants used fewer gestures when describing the rotation of spiky mugs (Chu & Kita, 2015). The fewer gestures reflected their avoidance of the spikes if required to physically rotate the cups.

The optimality principle of action identification theory states that over time and with repeated action, the person converges on a level that enables that individual to perform the action up to his or her capacity. However, non-optimal levels occasionally occur and not only

impair performance but have been shown to promote self-consciousness and anxiety (Vallacher, 2007).

A study on persuasion demonstrates the importance of finding an optimal level (Kim & Duhachek, 2020). Construal theory (Trope & Liberman, 2010) predicts that the most effective level of persuasion depends on how we perceive the intentions of the persuader. Although there had been research on persuasive messages provided by people, there has been a lack research on persuasive messages provided by nonhumans. Robots are programmed by humans to serve us, therefore, we should find their messages more persuasive when they inform us *how* to accomplish a task (a low-level action) than *why* we should accomplish a task (a high-level action).

A robot made a recommendation, supposedly based on a list of recently purchased products (Figure 6.12), but in fact the recommended product was randomly assigned at either a low-level or high-level of construal (Figure 6.13). Participants then rated the persuasiveness of the recommendation. As predicted, they found the low-construal level (Figure 6.13A) was more effective for a typical robot that would be unaware of higher-level explanations. However, the results were reversed when told the robot was self-aware of what is happening around it. Participants who received a recommendation from the informed robot rated the high-construal message (Figure 6.13B) as more persuasive. The optimal level for persuasion in this study depended on how participants perceived the self-awareness of the messenger.

### Neuroscience of Actions

The distinction between low-level and high-level actions indicates a distinction between mechanical actions (hitting a tennis ball) and executive actions (selecting a shot). Low-level actions depend on procedures; high-level actions depend on reasoning. Procedural memory supports low-level actions; declarative memory supports high-level actions.

The challenge is to integrate the contributions of procedural and declarative memory when both are needed. Osiurak and Heinke (2018) propose a theoretical framework shown in Figure 6.14. Tools used for manual work rely only on procedural memory because there are standard procedures for using these tools. In contrast, technical reasoning relies on using tools for non-standard procedures (free tool use) and therefore requires both procedural and declarative memory. Reasoning skills are also required when one performs a task for the first time.

Try to imagine a person who has never sliced a tomato and has to select one of the three tools shown in Figure 6.15 (Osiurak et al., 2020). The person can use mechanical declarative knowledge of cutting to select the knife as the most promising tool. Mental simulation of the action supports planning and evaluation before selecting motor actions to execute the procedure. If the action is unsuccessful, there is another attempt to select an appropriate tool. The right half of Figure 6.15 illustrates a neurocognitive perspective. A region (blue) in the inferior parietal lobe (IPL) is a likely candidate for integrating the mechanical actions in another region (purple) of the inferior parietal lobe (IPL) with motor actions (yellow) in the inferior parietal sulcus (IPS).

There is not only neurological evidence for the mental simulation of actions but support for the claim that, in some cases, the simulated actions can be identified from fMRI recordings (Mason & Just, 2020). The identification is aided by the fact that, unlike a declarative concept,

simulating actions occurs over the time period required to perform the action. The actions in this case consisted of learning to tie seven nautical knots, such as the one shown in Figure 6.16.

After learning to tie the knots, the adult participants imagined tying each knot while the investigators took fMRI recordings. A machine-learning algorithm then analyzed the recordings to identify areas in the brain that were the most diagnostic in identifying which knot was involved in the mental simulation. The ellipses in Figure 6.17 depict the most diagnostic regions. The regions (from left to right) are the language and executive area in the frontal brain, the motor cortex, the parietal area for spatial processing, and the cerebellum.

Mason and Just (2020) hypothesized that the parietal area likely supported the simulated actions while the motor cortex and cerebellum supported a motor plan for execution. They recommended that future studies should include tasks that are more conceptual, such as computer programming. This recommendation would be helpful in evaluating the technical reasoning framework developed by Osiurak and his colleagues (Osiurak et al., 2020; Osiurak & Heinke, 2018).

## APPLICATIONS

### Cognitive Offloading

Let's now return to the concept of cognitive load introduced in Chapter 4 by examining how it can be reduced by cognitive offloading. Cognitive offloading occurs when a physical action reduces the cognitive demands of the task (Risko & Gilbert, 2016). The physical action may either directly involve the body or involve the environment. For instance, rotate this text 45° to the right. You will likely find it easier to read if you also rotate your head 45° to the right. Another example of using the body to make a task easier is the use of fingers in counting by young children. Offloading information onto the environment occurs when we enter appointments into a calendar, create a shopping list, or post sticky notes (Photo 6.5).

As illustrated in Figure 6.18, cognitive offloading onto the environment requires making a decision when to use internal processes versus external aids (Risko & Gilbert, 2016). Evaluating our memory ability and the required effort to accomplish the task influence this decision. For instance, should a driver use unaided spatial memory or a GPS device to navigate to a friend's house? GPS is typically more reliable, but it may also prevent learning unaided navigation if we rely too much on it.

The overreliance on external aids was revealed in a study that provided participants with the opportunity to write numbers on a paper rather than simply remember them over a brief time period (Risko & Dunn, 2015). The research paradigm used a variant of a traditional short-term memory task that required recall of a string of auditorily presented digits. Participants knew the number of digits on each trial and, as expected, were more likely to write digits as the number increased from two to ten. They also wrote either all of the digits or none of the digits, so they would not have to decide which to record and which to remember. A surprising finding was that almost half of the participants wrote the digits even when told there would be only two digits.

Gestures use the body, rather than the environment, to offload information. Gestures are helpful during instruction because they enable teachers and students to demonstrate concepts such as size that may be difficult to state verbally. They can also facilitate reasoning by reducing the demands on working memory. Wagner et al. (2004) asked college-age adults to factor quadratic equations on a whiteboard and then explain their solutions. To determine the memory demands of the explanation task, the researchers gave the students supplementary information, such as a random string of letters, before their explanations. Students were later able to recall more supplementary items if they gestured while explaining their solution. Gesturing reduced the memory demands, particularly when the gestures and verbal explanations were compatible.

## Instruction

Instructors can aid learning by designing tasks that are supported by actions. Arthur Glenberg, a major contributor to theories of embodied cognition, and his co-investigators have tested the implications of the theory for helping young readers (Walker et al., 2017). An iPad app called *EMBRACE* (Enhanced Moved By Reading to Accelerate Comprehension of English) is designed to help English-language learners improve their comprehension of written English. It also helps them learn about science.

Figure 6.19 shows a page from a text used to teach early readers about Newton's three laws of motion. Many educators might think it too early to try to teach these concepts to children because they seem so abstract. But motion is something we experience all of the time. The key is to use embodied processing to teach children to map unfamiliar words, such as "force" and "acceleration," to actions and experiences that are familiar.

When Juan ties the small dog to the cart, children can use their sensorimotor experience to simulate the dog pulling the cart. But according to Newton's second law,  $\text{force} = \text{mass} \times \text{acceleration}$ , a larger force will be needed when Maria puts the logs into the cart, thereby increasing its mass. Now children can simulate the dog having a hard time providing enough force to get the cart moving—that is, to accelerate it. But, on the next page, Maria ties the horse to the cart, and the horse can apply enough force to easily move the cart.

Does this really work? Gómez and Glenberg (2020) reported on research conducted in Chile with children in the second and third grades reading the Spanish translation of the text. After reading each chapter, the children took a multiple-choice comprehension test. Children who acted out the text, either by moving the pictures on the iPad or by pantomiming, showed significantly greater comprehension compared to children who simply read the text. Also, children who acted out the text did significantly better on a vocabulary test that consisted of words such as "force" and "acceleration."

*EMBRACE* is an excellent example of a learning system that applies virtual actions to virtual objects. A big question for instructional design is how to choose between virtual and physical representations. To help answer this question, Martina Rau (2020) reviewed research on the effects of physical and virtual learning in science, technology, engineering, and mathematics. Her article investigates (1) what predictions different perspectives make about these two types of representations and (2) whether these predictions conflict or align with each other on issues such as physical engagement, cognitive load, and conceptual salience. For instance, Rau

suggests the physical interactions can be more engaging but virtual interactions are more effective in reducing cognitive load.

A second theme in Rau's (2020) article is the design of future research to evaluate the effectiveness of blended technologies, such as displaying the effects on a computer screen while manipulating physical objects. Such research needs to identify how different learning mechanisms interact with each other to determine how to productively create blended technologies.

## SUMMARY

The action-specific account of perception proposes that people perceive their surrounding environment in terms of their ability to act on it, ranging from daily activities to superior athletic performance. According to the perceptual symbol systems theory, people directly store sensory experiences in memory and can re-experience them through mental stimulation. Simulations that occur during language comprehension can result in gestures according to the gestures-as-simulation hypothesis. Gestures can then activate, manipulate, package, and explore spatial information during both speaking and thinking.

Physical, virtual, and mental actions can be combined with physical, virtual, and mental objects. Examples of physical actions on physical objects are Montessori manipulatives, on virtual objects are Wii sports games, and on mental objects are gestures. Examples of virtual actions on physical objects are robotic surgery, on virtual objects are virtual experiments, and on mental objects are teaching the blind. Examples of mental actions on physical objects are brain-computer robotic interfaces, on virtual objects are brain-computer cursor interfaces, and on mental objects are sports simulations.

High-level actions, such as holding, getting, and pouring milk, depend on low-level actions, such as movement of the joints and muscles. Modifying actions for activities as diverse as typing and high jumping requires discovering new components that initially degrade performance but can eventually result in substantial improvements. Action identification theory proposes that action is maintained with respect to its available identity, there is a bias toward a higher-level identity, a shift to a lower-level identity when necessary, and a tendency to find an optimal level of performance. Neuroscience offers support for a theoretical framework in which technical and novel actions rely on both procedural and declarative memory.

Cognitive offloading occurs when a physical action reduces the cognitive demands of the task. The physical action may either directly involve the body (such as gesturing) or incorporate the environment (such as making a list). An iPad application called EMBRACE (Enhanced Moved By Reading to Accelerate Comprehension of English) helps English-language learners improve their comprehension of written English and learn about science. The key is to use embodied processing to teach children to map unfamiliar words, such as "force" and "acceleration," to familiar actions and experiences. Future research needs to evaluate the effectiveness of blended technologies on instruction to identify how different learning mechanisms interact with each other.

**RECOMMENDED READING**

Action has become an increasingly important contributor to understanding the cognitive and neuropsychological aspects of cognition (Engel et al., 2013). Exploration is a unifying construct that applies across many domains, ranging from animal foraging to cultural innovation (Hills et al, 2015). Exercise, sports, and the performance arts link skill training with cognition (Tompsonowski & Pesce, 2019). Knowledge of actions is useful for how we perform, simulate the actions of others, and acquire a conceptual understanding of performance (Quandt & Chatterjee, 2015). Risko and Gilbert (2016) discuss the mechanisms that trigger cognitive offloading and the consequences of this behavior.