Blackwell Companions to Philosophy

This outstanding student reference series offers a comprehensive survey of philosophy as a whole. Written by today's leading philosophers, each volume provides lucid and engaging coverage of the key figures, terms, topics, and problems of the field. Taken together, they provide the ideal basis for course use, representing an unparalleled work of reference for students and specialists alike.

Published

1 A Companion to Ethics
Edited by Peter Singer

2 A Companion to Aesthetics
Edited by David Cooper

3 A Companion to Epistemology
Edited by Jonathan Dancy and Ernest Sosa

4 A Companion to Contemporary Political Philosophy
Edited by Robert E. Goodin and Philip Pettit

5 A Companion to the Philosophy of Mind
Edited by Samuel Guttenplan

6 A Companion to Metaphysics
Edited by Jaegwon Kim and Ernest Sosa

7 A Companion to Philosophy of Law and Legal Theory
Edited by Dennis Patterson

8 A Companion to Philosophy of Religion
Edited by Philip Quinn and Charles Taliaferro

9 A Companion to the Philosophy of Language
Edited by Bob Hale and Crispin Wright

10 A Companion to World Philosophies
Edited by Eliot Deutsch and Ron Bontekoe

11 A Companion to Continental Philosophy
Edited by Simon Critchley and William Schroeder

12 A Companion to Feminist Philosophy
Edited by Alison Jaggar and Iris Young

13 A Companion to Cognitive Science
Edited by William Bechtel and George Graham

---

*Blackwell
Companions to
Philosophy*

A Companion to Cognitive Science

Edited by

WILLIAM BECHTEL and
GEORGE GRAHAM

Advisory editors

David A. Balota
Paul G. Chapin
Michael J. Friedlander
Janet L. Kolodner

BLACKWELL
Publishers

First published 1998

2 4 6 8 10 9 7 5 3 1

Blackwell Publishers Inc.
350 Main Street
Malden, Massachusetts 02148
USA

Blackwell Publishers Ltd
108 Cowley Road
Oxford OX4 1JF
UK

Library of Congress Cataloging-in-Publication Data

A companion to cognitive science / edited by William Bechtel and
George Graham; advisory editors, David A. Balota . . . [et al.].
p. cm. — (Blackwell companions to philosophy; 13)
Includes bibliographical references and indexes.
ISBN 1-55786-542-6 (hardcover: alk. paper)
1. Cognitive science. I. Bechtel, William. II. Graham, George,
1945– . III. Balota, D. A. IV. Series.
BF311.C578 1998
153—dc21 97-38757
CIP

British Library Cataloguing in Publication Data

A CIP catalogue record for this book is available from the British Library.

Typeset in 10 on 12 pt Photina
by Graphicraft Typesetters Limited, Hong Kong
Printed in Great Britain by T.J. International, Padstow, Cornwall

This book is printed on acid-free paper

Contents

PART I

THE LIFE OF COGNITIVE SCIENCE

WILLIAM BECHTEL, ADELE ABRAHAMSEN,
AND GEORGE GRAHAM

Contents

Preliminaries

Let's begin prematurely. Let's try to characterize cognitive science:

> Cognitive science is the multidisciplinary scientific study of cognition and its role in intelligent agency. It examines what cognition is, what it does, and how it works.

That proposition may appear more definitive than it truly is. Which creatures or sorts of things count as intelligent agents? Insofar as cognitive science seeks to be multidisciplinary, which scientific disciplines are included? Do they interact substantively – share theses, methods, views – or do they simply converse? Finally, how does one discover what cognition is, what it does, and how it works? Cognitive scientists answer these questions in a variety of ways. No answer is without dissent. Each inspires controversy: everyone likes some answer, but no one likes every answer.

Shall we chart the answers? Only a conceptual botanist would delight in that task; besides which, it would be a premature and unhelpfully abstruse way in which to introduce both cognitive science and the content of this Companion. To those two related ends we prefer a short anecdote, then a long story – a very long story. We shall revisit the above characterization at the very end of the story, for by then the abstruse will have metamorphosed into the familiar, and any sources of controversy will be intelligible if not eliminable.

An anecdote: Building 20

Though all three of us objected to the Vietnam War, one of us (GG) was formally classified as a conscientious objector and, during the early 1970s, performed civilian alternative work service for New England Deaconess Hospital in Boston. One day – his day off – on a rather aimless walk through the campus of Massachusetts Institute of Technology in Cambridge, he came upon some stoically wooden buildings set unobtrusively in the middle of the campus. One was marked simply "Building 20." Looking for a telephone, the future co-editor asked a student standing in front of the building, "Is there a public phone in 20?" "I don't know," replied the student. "All I know about 20 is that Noam Chomsky works here."

"Noam Chomsky?" One hates to admit such ignorance, but being new to Cambridge and unfamiliar with *Syntactic Structures*, perhaps one can be forgiven.

"What?!" Befuddled, but trying to be polite: "Why, he's the world's leading linguist." In retrospect, I had stumbled into the domain of one of the prime movers of modern cognitive science. Chomsky was both icon of the Cambridge anti-war movement and hero of the battle against anti-cognitive psychology – behaviorism.

"Without Chomsky," added the student, "you would be left with B. F. Skinner and his rats up at Harvard."

It was the early 1970s. Talk of cognition thickened the air; cognitive science was growing up. So how *did* cognitive science form? How did it self-conceive and mature? Certainly Chomsky played a key role. Others did too. Time for the long story.

A predecessor: behaviorism

In North America something dramatic happened in psychological science in the 1950s, something often referred to, in retrospect, as the *cognitive revolution*, something Howard Gardner characterized as "the unofficial launching of cognitive science" (Gardner, 1985, p. 7). The revolt was against behaviorism, which was heralded in John Watson's 1913 manifesto and quickly came to largely dominate psychology and linguistics, and influence other disciplines in North America. Behaviorism turned away from earlier, mentalistic attempts to analyze the mind; instead it focused on overt behavior and the discovery of regularities involving observable events and behaviors. "Psychology," wrote Watson, "as the behaviorist views it is a purely objective experimental branch of natural science" (1913, p. 158). Behaviorism was a blend of Darwinism, functionalism in psychology, and anti-introspectionism. It was a normative meta-psychology; it tried, from its own platform, to legislate psychologists into being good empirical scientists. Here, very quickly, most roughly, and simplified stepwise, is how behaviorism said psychology should be done:

> Step One: Observe behavior.
>
> Step Two: Select descriptions of behavior which are nonmentalistic – that is, which do not presuppose theorizing about the internal psychology of the organism or agent in question.
>
> Step Three: Select descriptions of the environment (in which the observed behavior takes place) which themselves are nonmental in that they do not presuppose theorizing about how the organism or agent represents its environment.
>
> Step Four: Note that certain nonmental aspects of behavior (such as its frequency of occurrence, physical direction, and so forth) seem to be correlated with certain nonmental aspects of the environment (physical stimuli which are present).
>
> Step Five: Judiciously vary – in a laboratory model and experimental setting – the environmental aspects; thereby determine the class of environmental events and the class of behaviors covered by the correlation.

> Step Six: Speak of the behavior (response) as a function of the environment (stimuli); refer to environmental stimuli and behavioral responses as existing in a functional relationship.

A compressed example illustrates:

> A rat scurries across the alley. It turns left towards a tipped garbage can and ingests food. Remove the rat from the alley. Place it in a laboratory maze. Vary the location of food pellets with the direction of its turning (whether it turns left or right). Note that under certain conditions the behavior of turning left or right is correlated with its immediate history of ingesting food. The history is "responsible" for the direction. Left turning is a function of a food-left history; right turning is a function of a food-right history.

The specification of functionally related stimuli and responses posed a number of problems for behavioristically oriented psychology, itself sometimes called "the experimental analysis of behavior." Often, for example, stimuli and responses selected for a functional class cannot be usefully characterized in an *apsychological* (nonmental) vocabulary. Consider, for example, the temptation to classify the rat's responses as *seeking* food and *remembering* whether it was found to the left or right. Mentalistic attribution is a tough temptation to resist. In some cases – human verbal behavior, for instance – it is impossible to resist. However, let's return to the chronology.

In North America behaviorism reigned for decades as a remarkably resilient, influential, and in many ways laudable doctrine that resonated through a number of disciplines beyond psychology. In linguistics it helped to displace philology (the study of the histories of particular languages) with empirical studies of language use. Under the leadership of Leonard Bloomfield, linguistic behaviorism aspired to carry out a program in which linguists would collect speakers' utterances into a corpus and produce a grammar that described it. Explicitly excluded were any mentalistic assumptions, inferences, or explanations.

In philosophy, the logical positivism of Rudolf Carnap and Carl Hempel was congenial to behaviorism. Each tried to develop behavioristic canons for the meaningfulness and empirical grounding of scientific hypotheses. Hempel himself eventually abandoned this effort: "In order to characterize the behavioral patterns, propensities, or capacities . . . we need not only a suitable behavioristic vocabulary, but psychological terms as well" (Hempel, 1966, p. 110). Others maintained a thoroughgoing empiricism. Willard van Orman Quine imposed behavioristic standards on the task of interpreting the speech of another person (or oneself) and argued that the only evidence available was the sensory input from the environment. He argued that from this evidence alone the meaning of a sentence would always be indeterminate, and therefore concluded that the notion of meaning was vacuous. He made an exception only for those statements most firmly rooted in sensory experience (observation statements).

The story to be told

Not everyone agreed with behaviorist strictures. To such critics as the aforementioned resident of Building 20, behaviorism was a severely truncated, virtually atheoretical stance. The historical events to be discussed in the next section clearly represent a rebellion against behaviorism and the birth of a new approach. The first stirrings of life of a cognitive science revolution occurred at the end of World War II. The concept of *information* came to center stage in cybernetics, information theory, and early neural networks. This enabled cognitive researchers to cast off their fear of mentalism and attempt to understand the processing of information in the bead – in the mind – that underlies behavior. By the mid-1970s the conceptual and methodological frameworks of linguistics, psychology, and philosophy were fundamentally altered in ways characteristic of what Thomas Kuhn (1962/1970) has referred to as a "scientific revolution." A generation of new thinkers, including Chomsky, George Miller, and Hilary Putnam, had created a new *paradigm*, and a new generation of researchers took up the banner and gave birth to a radically different set of research agendas. In addition, a brand new discipline – artificial intelligence – emerged, and such leaders as Allen Newell and Herbert Simon linked its approach to those of the other disciplines. Neuroscience also made major advances, but within its own paradigm.

The story that follows is about the development of cognitive science as an intellectual enterprise *and* as an institution. The choice of conjunction is intentional. The intellectual enterprise of cognitive science did not develop independently of its institutions; so if we are interested in the enterprise of cognitive science, we need to mention the kinds of social mechanisms, ranging from journals to graduate programs, that often both reflected and helped to support intellectual changes.

The story that we tell, like all stories, is selective. Without filling the entire volume with historical narrative, there is no way we could cover all the plot lines of research and theory that contributed to what is now known as cognitive science. We have chosen to emphasize work that is interdisciplinary in its nature or impact, with the result that a number of major researchers doing core work within each discipline have been left out. Our other constraint is our goal of providing a context for the contemporary scene, especially as it is portrayed in the rest of the volume. Accordingly, this essay is skewed toward earlier developments, becoming briefer as we get closer to the current scene. The reader is encouraged to follow the links to the other contributions, which we have marked by setting their titles in small capitals. These links become ever more numerous as we go along, so as to fill out the story.

1 Gestation and birth of the cognitive revolution

Cognitive science did not emerge suddenly. Like a person, it went through a long period of gestation. Unlike a person, it has no official and unambiguous birthdate. Following one of its pioneers, George Miller, we have selected 1956 as a plausible year of birth. Most of the preparatory developments occurred in computer science, psychology, and neuroscience. Only in the last stages of this gestation did Chomsky arrive on the scene to begin transforming linguistics.

1.1 *The seeds of computation*

The attempt to design intelligent machines played a critical role in the development of cognitive science. Three different research traditions contributed to the development of such machines: cybernetics, artificial neural networks, and symbolic artificial intelligence (AI). While symbolic AI garnered attention as a central contributor when cognitive science took formal, institutional shape in the 1970s, both cybernetics and artificial neural network research played a major historical role as well, providing many of the ideas that allowed for characterization of events inside a person's head in cognitive or information processing terms. Early developments in artificial neural networks will be discussed in the context of neuroscience in section 1.3.5. For now, let's focus on cybernetics and artificial intelligence.

1.1.1 *Cybernetics*

How do living organisms maintain themselves in the face of changing and often threatening external environments? Here, roughly, is the answer of Claude Bernard, a mid-nineteenth-century physiologist: "Each living organism is composed of different component sub-systems; these respond when particular features of the organism's internal environment exceed – under pressure from its external environment or malfunctioning of a sub-system – their normal range. They act so as to restore that feature to the normal range." Bernard's notion of internal componential adjustment to external change contained the germ of the notion of feedback, which is central to cybernetics.

The idea of feedback was developed more thoroughly by Norbert Wiener, who had interests and training in biology before getting a degree in mathematical logic. He conceptualized feedback as consisting in the feeding of information generated by a system back into the system, thereby enabling it to adjust its behavior. Wiener (1948) coined the term *cybernetics*, which derived from the Greek for *helmsperson*, for this idea, and he proposed that natural and artificial systems could steer themselves by using feedback.

At MIT, where he was professor of mathematics, Wiener collaborated with Vannevar Bush, who in the 1930s had begun to develop an analog computer. At the start of World War II, they set out to design a system for improving anti-aircraft fire in which feedback would play a critical role. Information from radar would be employed to calculate adjustments to gun controls; after new shots were fired, information about the results would be used to readjust the gun controls. If this rather Bernardian procedure were automated, one would have a self-steering device; even if humans were part of the loop, the overall activity would count as one of self-steering by means of feedback.

As the war continued, Wiener collaborated with two other researchers, Julian Bigelow (an engineer) and Arturo Rosenblueth (a physiologist). The three scientists offered a cybernetic theory of "control and communication in the animal and machine." Rosenblueth gave the first public presentation in 1942 at a conference on Cerebral Inhibition sponsored by the Josiah Macy Foundation, which was soon to play a critical role in developing the cybernetic framework. Together Rosenblueth, Wiener, and Bigelow published a paper entitled "Behavior, Purpose and Teleology" in *Philosophy of Science* in 1943, in which they ventured to use the concept of feedback to legitimize the notion of *teleology* (goal direction); in their view, feedback enabled systems, both living and artificial, to be goal-directed.

In January 1945, Wiener, together with Howard Aiken and John von Neumann, brought together a group of theorists from a broad range of disciplines for a meeting on the notion of feedback. Among the participants were the neurophysiologists Warren McCulloch and Rafael Lorente de Nó, the logician Walter Pitts, and Samuel Wilkes (a statistician), Ernest Vestine (a geophysicist), and Walter E. Deming (a quality control theorist). On January 24 Wiener wrote to Rosenblueth (who had returned to Mexico), emphasizing the potential for integrating the study of the brain with engineering work on artificial systems:

> The first day von Neumann spoke on computing machines and I spoke on communication engineering. The second day Lorente de Nó and McCulloch joined forces for a very convincing presentation on the present status of the problem of organization of the brain. In the end we were all convinced that the subject embracing both the engineering and the neurology aspects is essentially one, and we should go ahead with plans to embody these ideas in a permanent program of research. (Quoted in Heims 1980, p. 186)

In addition, the participants talked of creating a journal and a scientific society after the war. These plans did not come to fruition, but Heinz von Foerster and McCulloch, with support from the Macy Foundation, did organize twice-yearly meetings of the group and invited investigators from an even broader array of backgrounds, such as psychologist Kurt Lewin and anthropologists Gregory Bateson and Margaret Mead. Originally, the conference series was called the Conference for Circular Causal and Feedback Mechanisms in Biological and Social Systems, but in 1949 it adopted Wiener's term *cybernetics* and changed its name to the Conference on Cybernetics.

The last meeting was in 1953, and the cybernetics movement waned. Some of its key ideas, such as the notion of feedback of information from environment to behaving system, were further developed later by cognitive scientists. Cybernetics also represented a first attempt at a broad, multidisciplinary endeavor to explain mental phenomena. An especially noteworthy difference from the cognitive science of the 1970s was the central role of neuroscience in cybernetics. Some of its products, such as W. Ross Ashby's *Design for a Brain*, put forth ideas that were in a sense ahead of their time and are only now bearing fruition in cognitive science.

1.1.2 *Computers and artificial intelligence*

Of all the research fields that would come to play a major role in cognitive science, ARTIFICIAL INTELLIGENCE, usually classified as a branch of computer science, was the newest, having to await the invention of the computer itself. The digital computer, as we know it, was another product of World War II, though the idea of automated computing goes back much further. One key element of computing is the idea of a set of instructions that can be applied mechanically. An early version of this idea was found in an 1805 device of Joseph-Marie Jacquard which used removable punch cards to determine the pattern which a loom would weave. In the 1840s, Charles Babbage made use of this idea in his design for an *analytical engine*, which was to have been a steam-driven computational device. Babbage never succeeded in actually building the engine. He did, however, engage in a fruitful collaboration with Lady Lovelace (Ada Augusta Byron), who worked out ideas for programming Babbage's machine. Ada, the modern programming language, was named in her honor.

A major hurdle faced by Babbage in the nineteenth century was the lack of sufficiently precise manufacturing for the components of his engine. However, by the start of the twentieth century, precision had improved to the point where mechanical calculators could be manufactured by companies such as Tabulating Machine Company, which later merged into IBM. These machines were purely mechanical – without electrical components – but in the late 1930s Claude Shannon showed that electric switches could be arranged to turn one another on and off in such a way as to perform arithmetic operations. The idea of using electronic circuits to carry out calculations was put into practical use during World War II in England by Alan Turing and his collaborators at Bletchley Park in the effort to decipher German military communications. The German cipher machine Enigma was a particular challenge, since it was built out of a set of rotors which permuted the letters of the alphabet; the rotors were mechanically coupled so as to constantly change the alphabetic substitutions employed in the cipher. The challenge to Turing and his colleagues was to examine all combinations of encoding assignments in the machine to find the one used in the cipher – a huge computational task. The result was Bombe, which employed a single electronic valve for fast switching. For highest-level communications, Germany employed an even more sophisticated cipher, which produced what researchers at Bletchley Park referred to as "Fish" cipher text. To decipher these messages, Turing and his colleagues designed a vacuum tube-based special-purpose machine, Colossus, which employed thousands of electronic valves.

Another World War II era computer, ENIAC (Electronic Numerical Integrator and Calculator), was developed by J. Presper Eckert and John Mauchly at the Moore School of the University of Pennsylvania. It was designed to calculate artillery tables, which would specify how to aim artillery on various terrains so as to hit desired targets. Despite massive effort, ENIAC remained incomplete until 1946. John von Neumann designed the basic architecture for ENIAC – the "von Neumann architecture." It was, however, only fully realized in ENIAC's successor, EDVAC (Electronic Discrete Variable Computer), and has continued to play a central role in computing to the present.

At the heart of von Neumann architecture is a distinction between a computer's memory and its central processing unit (CPU). One of von Neumann's innovations was to recognize that the instructions comprising a program could be stored in memory in the same manner as the data being operated upon. Computer operations are carried out in cycles in the CPU; in each cycle both data and instructions are read from memory into the CPU, which carries out the instructions and returns the results to memory.

We now come closer to the role of the computer in the birth of cognitive science, but we need to make another brief digression. After the war, computers became increasingly powerful. And with such power a possibility began to be realized that had first been envisaged by Gottfried Wilhelm Leibniz, the famous seventeenth-century philosopher at the University of Leipzig. He had proposed that numbers could be assigned to concepts, and that formal rules for manipulating those numbers would in effect also manipulate the concepts to which they were assigned. In 1854, the English mathematician George Boole had taken a major step in developing this idea in a book called *The Laws of Thought*. Boole formulated several operations that could be performed on sets. He also showed that these operations correspond to logical operators (*and*, *or*, *not*) which could be applied to propositions. He suggested that the laws governing these operations could serve as laws of thought. The switches that Shannon had devised in the late 1930s performed these basic Boolean operations, with the resulting state of the switches (*on* or *off*) corresponding to the truth values of the proposition (*true* or *false*).

Boole's system was limited to operations on complete propositions (e.g., "The woman is a lawyer") and could not deal with structure internal to the proposition (e.g., the fact that the predicate "is a lawyer" is being predicated of "the woman"). Gottlob

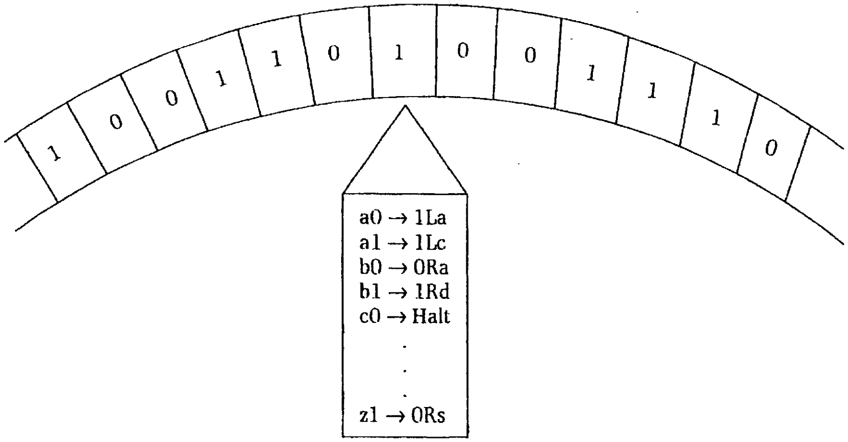

Figure I.1 A Turing machine.

Frege, though, expanded the system in 1879 to deal with such predications (permitting representations of arguments from premises such as "All lawyers have passed the bar exam" and "The woman is a lawyer" to "The woman has passed the bar exam"); the resulting system of predicate calculus provided a way of formalizing inferences that has been extremely influential. The idea of formally representing information in symbolic notation and using formal operations to transform this information provided a critical entrée to the use of computers to simulate reasoning.

Two additional ideas that have guided the use of computers to model thinking were contributed by Alan Turing. Even before the first digital computer was built, Turing (1936) proposed a simple machine for performing computations. The Turing machine (figure I.1) consisted of a read/write head and an infinite tape. The tape consisted of squares, each containing either a 0 or 1. The head could move one square to the left or right along the tape, read the numeral off that square, and write a replacement numeral. The head could be in any one of a finite number of states, each of which would specify what to do in response to a given number on the tape (e.g., if the square contains a 0, write a 1, move one square to the left, and remain in state *c*). Turing showed that for any well-defined series of formal operations (such as those of arithmetic), one could design a Turing machine which could carry it out. Also in the 1930s Alonzo Church independently proposed that any process for which there is a decision procedure could be carried out through such a series of operations; accordingly, the Church–Turing thesis proposes that any decidable process can be implemented in a Turing machine. Turing also proved that it was possible to construct a universal Turing machine that could simulate the operation of any given Turing machine; a universal Turing machine would thereby be capable of carrying out any well-defined series of operations.

If it were possible to provide it with infinite memory, a properly programmed von Neumann computer would be a universal Turing machine. The challenge would then be to provide it with the right program for carrying out all of the necessary formal operations. But is carrying out formal operations sufficient for thinking – for conceptual thought? Part of the difficulty in answering this question is that we lack a notion of thinking that is sufficiently clear for us to decide whether it could be accomplished in formal operations. What is needed is a way of specifying when something is thinking.

Turing had an ingenious proposal here as well. He offered a test – not the sole test, but a test – for thinking (Turing, 1950). His suggestion was to approach the question in terms of the behavior of the machine: could its behavior pass for that of a thinking person? If yes, it thinks. In what is now known as the Turing test, one decides whether a machine is thinking by arranging for a human interrogator at a keyboard to ask questions of both the machine and another human – both of whom are unseen but whose answers are displayed. If the interrogator, even after sophisticated questioning, cannot differentiate the computer from the human, then the computer's activity counts as thinking. Turing recognized that it would require a very complicated machine to engage in any protracted dialogue with humans and not be detected, but he believed that a computer would eventually pass this test.

By the early 1950s the theoretical foundations for artificial intelligence had been established; what remained was to actually build systems that exemplified aspects of human thinking. This task fell to a younger generation of investigators who were just then launching their careers. The team that set the prototype for the new enterprise of modeling intelligence – of producing an artificially intelligent system – consisted of Herbert Simon and Allen Newell. Neither Simon nor Newell was initially oriented towards computer science. Simon's background was in political science, and his appointment was in the Graduate School of Industrial Administration at Carnegie Tech (now Carnegie–Mellon University); he first made his reputation, and later won a Nobel Prize, in economics for his analysis of the functioning of human organizations. This work led him to challenge one of the tenets of modern economics, the assumption that agents are perfectly rational in the choices which they make (see Article 57, INSTITUTIONS AND ECONOMICS). Simon, to the contrary, emphasized that rationality was *bounded* and that, rather than examining all possibilities they face and then choosing one, humans generally accept the first option which meets a predetermined standard. Simon called this approach to decision making *satisficing* (see Article 44, HEURISTICS AND SATISFICING). He also drew from his work on human organizations the recognition that humans often rely on stock recipes, or *heuristics*, rather than seeking optimal solution procedures that guarantee correct answers. Lastly, Simon noted how lower divisions of a corporation typically pursue subgoals of the corporation's overall goal, thus suggesting the strategy of subgoaling in computer programs.

Starting in 1952, Simon became a consultant for the RAND Corporation, and on a visit to the RAND offices was intrigued by a printer he saw producing maps using characters other than numerals. He found the idea of a computer manipulating non-numeric symbols attractive. At RAND he also met Allen Newell, who was developing such maps as part of a project of modeling an air defense center. The prospect that intrigued Newell came from Oliver Selfridge, who was using digital computers to simulate a neural net-like model for pattern recognition which he dubbed "Pandemonium." Such a model organizes a number of specialized agents (*demons*) into layers, with those in each layer competing in parallel to recognize a pattern. The output of lower layers serves as input to higher layers, as illustrated in figure I.2 for a model of letter recognition. Selfridge's work led Newell to the notion of a complex process being achieved through the interaction of simpler subprocesses.

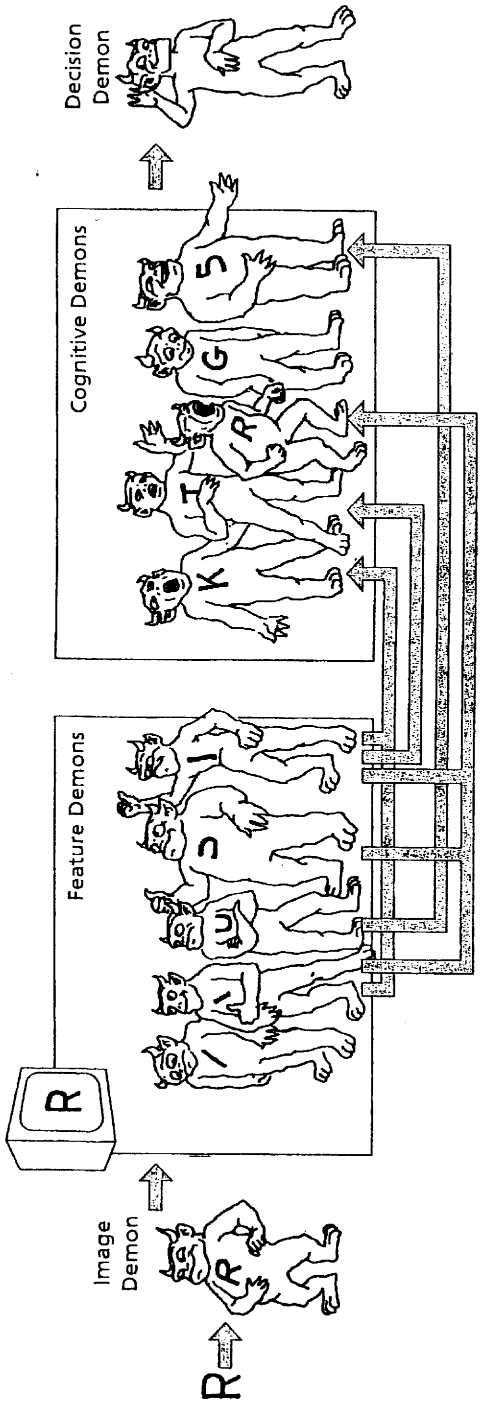

Figure 1.2 Rendition of Oliver Selfridge's Pandemonium model. The *demons* at each level beyond the image demon (which merely records the incoming image of a letter) respond to those demons at the preceding level with which they have links (indicated by arrows). Figure drawn by Jesse Prinz.

Having both been drawn to ideas about problem solving through means–goal (means–end) reasoning via heuristics, Simon and Newell set out with computer programmer J. Clifford Shaw to develop a system that could prove theorems in formal logic. They first implemented the system with human agents (including Simon's children) playing the roles of the various parts of the program; each agent carried out rules written on index cards. When actually programmed in 1956, their so-called Logic Theorist proved 38 theorems from Russell and Whitehead's *Principia Mathematica*, one more elegantly than had Russell and Whitehead. The actual implementation of Logic Theorist represented more than the first apparent success of a computer program in performing a task requiring intelligence; it also brought the development of a list processing language, IPL (Information Processing Language). The symbols in a list could be stored at arbitrary memory addresses in the computer, and links could be added between one item and another simply by specifying at the first site the address of the related item.

Two other important pioneers in artificial intelligence were Marvin Minsky and John McCarthy. Minsky, after completing a dissertation on neural networks at Princeton in 1954, was drawn to modeling intelligence by writing programs for von Neumann-style computers. McCarthy also did his doctoral work at Princeton, first doing research on finite automata (systems like the read/write head of a Turing machine which could, by following rules, progress through a finite number of different states). During a summer at IBM in 1955, though, he too became attracted to modeling intelligent processes on digital computers. Minsky and McCarthy joined forces to organize a pivotal two-month workshop, the Dartmouth Summer Research Project on Artificial Intelligence. They secured funding from the Rockefeller Foundation, which enabled eight key researchers to join them at Dartmouth College during the summer of 1956. Nathaniel Rochester and Oliver Selfridge were using digital computers to simulate neural networks. The organizers and most of the other participants (Ray Solomonoff, Trenchard More, and Arthur Samuel, as well as Simon and Newell) were taking advantage of the computer's ability to manipulate symbols to simulate thinking more directly. The research of the eighth participant, Claude Shannon, involved information theory (see below) and automata theory. Most of the discussion during the conference was programmatic, taking off from the proposition initially put forward in the application to the Rockefeller Foundation that "every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it." In some respects, the highlight of the conference was the presentation of Logic Theorist, the prototype of an intelligent system. However, this also became the source of some social tension; because Newell and Simon were busy programming Logic Theorist during the summer, they attended the conference for only one week.

At the Dartmouth Conference the new enterprise was christened *artificial intelligence (AI)*. However, as so often happens at christenings, the name served both to unify and to divide. *Artificial* suggested that the form of intelligence exhibited by computers might differ from that of humans, and indeed, at the time, Minsky and McCarthy were not strongly committed to the idea that artificial intelligence would be particularly revealing about human cognition. Newell and Simon, who had the only working program, were more concerned with human cognition and were more directed toward psychology. For a number of years they resisted the label "artificial intelligence" and referred to "complex information processing." The very idea of information processing has an ecumenical conceptual feel to it, suggesting that just as humans process information, so, too, could computers, and perhaps even in the same way.

1.2 *Quickening: psychology makes its moves*

Many of the developments prior to the actual birth of cognitive science – developments discussed above – involved mathematicians and engineers laying the foundations for artificial intelligence. But AI was not the natural disciplinary locus for the study of thinking and reasoning. That, obviously, had to be psychology. And indeed, as cognitive science took life, psychology was one of its most important disciplinary contributors.

1.2.1 *Origins of psychology*

There is a very long history of inquiry into psychological questions within philosophy, and this left a substantial endowment to the newer discipline of psychology that emerged in the late nineteenth century; the best-known example is the idea that mental life can be understood in terms of elementary sensations that are combined by association (*associationism*). Nonetheless, it was developments in nineteenth-century German physiology that gave the immediate impetus to the emergence of psychology as a distinct, experimental science of the mind. Between 1850 and 1870, for example, Hermann Helmholtz devised a way to measure the speed of nerve transmission, proposed the notion of *unconscious inference* for operations occurring in the mind outside the reach of consciousness, and made major contributions to understanding how we see and hear (including his component theory of color vision and his studies of how subjects reorganize their perceptual experience after wearing distorting lenses for some time).

Helmholtz preferred not to label his work as psychology, but his laboratory assistant from 1858 to 1864 in Heidelberg, Wilhelm Wundt, was calling his own course "Physiological Psychology" by 1867 and had completed a major book on this topic by 1874. Moving to the University of Leipzig in 1875, Wundt established a demonstration laboratory which he upgraded to a research laboratory in 1879 – the event most often cited as the birth of psychology. Wundt's reach was extraordinary: he pursued more topics using a greater diversity of methods, taught more students (28,000), and published more pages (almost 60,000) than any of his friends or foes. During the last 20 years of his career, he focused on the use of nonexperimental methods to understand social life and language use. In the *Völkerpsychologie* that resulted, Wundt stretched beyond his inherited tradition of associationism to emphasize sentences as structured wholes that could not be reduced to a succession of words. Prior to this, Wundt was an experimentalist. In one line of research, he refined the Helmholtz–Donders technique of mental chronometry to produce surprisingly modern reaction time studies (see Article 27, BEHAVIORAL EXPERIMENTATION). He obtained time estimates for processes beyond simple perception, including *apperception* (achieving a synthesis or awareness of structure by means of conscious attention to a particular stimulus), cognition (required to discriminate one stimulus from others), and association (required when different stimuli have different responses). Wundt is better known, however, for his use of *introspection*, in which highly trained observers systematically analyzed their own mental experiences in order to identify the elements. Edward Titchener, the leading expositor of Wundt in the English-speaking world, focused almost exclusively on this technique.

Introspection was a major target of behaviorists; accordingly, Wundt's legacy of opening up a variety of scientific approaches to the study of mental activity was largely overlooked in North America during and after the ascendancy of behaviorism.

Whereas Wundt developed experimental foundations for the study of mental experience, William James emphasized the theoretical. At Harvard he too created a demonstration laboratory for psychology in 1874–5, but his genius lay in synthesizing and interpreting others' research rather than carrying out detailed experimental manipulations and measurements himself. Influenced by Darwin, James emphasized the adaptive function of behavior. His functionalism and his concern with individual differences contrasted with Wundt's structuralism and Kantian emphasis on intrapsychic universals – a theoretical divide that is still with us today in the camps of language acquisition researchers who emphasize language function and conversation versus those who follow Chomsky, for example. James developed his ideas in a monumental and engagingly written work, *Principles of Psychology*, ten years in the writing and finally published in 1890. His phenomenological descriptions of how mental activity is consciously experienced and how it figures in ordinary life are unparalleled and still frequently quoted. For James, habit functions as "the enormous fly-wheel of society, its most precious conservative agent" (vol. 1, p. 121), and consciousness is a continuously changing "stream of thought" (ibid., p. 224) rather than a construction of elements. James also contributed an influential theory of EMOTIONS, in which he claimed that "we feel sorry because we cry" (vol. 2, p. 450) rather than the reverse. By contrast, Wundt painstakingly gathered and analyzed introspections, arriving at three dimensions of variation in emotion: pleasantness/unpleasantness, activity/passivity, and tension/relaxation – which are strikingly similar to the results of Osgood, Suci, and Tannenbaum's (1957) factor analysis of rating scale data in the 1950s (section 1.2.2).

From its beginnings with Wundt and James, psychology developed quickly as a discipline. Many universities created chairs or departments of psychology, and the American Psychological Association was founded in 1892. Psychology was thus well established prior to the emergence of behaviorism, which not only repudiated the early tool of introspection but also, in the radical form advocated by John Watson in his manifesto noted above, prohibited any appeal to mental processes in explaining behavior.

1.2.2 *The era of behaviorism*

Watson can be credited with founding the behaviorist movement in the United States, but there are many other portraits with many other names in the behaviorist gallery: among others, Ivan Sechenev, Ivan Pavlov, Edwin Guthrie, Edward Thorndike, Edward Tolman, Clark Hull, Kenneth Spence, and, of course, B. F. Skinner. Some of these behaviorists were actually interested in cognition. Tolman, in particular, studied the ability of rats to navigate their environments and proposed that they did so by constructing *cognitive maps*; he also demonstrated learning without reward (*latent learning*) and posited a role for expectations and other *intervening variables*. He tried to show, in the words of one observer, "that a sophisticated behaviorism can be cognizant of all the richness and variety of psychological events" (cited in Bry, 1975, p. 59).

Tolman's primary rival, neobehaviorist Clark Hull at Yale University, pursued a less cognitive theory of stimulus–response learning. In his enormously influential 1943 book, *Principles of Behavior*, Hull systematized many of the research findings on instrumental conditioning in a *mathematico-deductive* theory. For Hull, how quickly a rat responded by pressing a lever (a measure of response strength) depended on such input variables as hours of food deprivation and number of trials reinforced by a food pellet. These are all observable variables; the tricky part was arriving at equations that would link them to posited intervening variables. *Drive* was a function of number of hours of food deprivation; *habit strength* was a negatively accelerated function of number of reinforcements and of drive reduction; and the *excitatory potential* that would lead to the actual response was a multiplicative function of drive, habit strength, and other variables. Hull's theory set much of the research agenda on learning in the 1940s. All too often, research results failed to support the theory, and it would be rescued with revisions until the next challenge.

Hull's camp (including Kenneth W. Spence at Iowa) continued sparring with Tolman's camp, but by the time of Hull's death in 1952 it was becoming increasingly obvious that neither theory had won and that no new grand theory would emerge. Tolman's cognition-friendly approach would not be appreciated again until the late 1960s. For those students who still wished to study animal learning, B. F. Skinner's (1938) *operant conditioning* paradigm was waiting in the wings at Harvard and was dominant by the late 1950s. Skinner designed elaborate schedules of reinforcement that produced pleasing regularities in the timing of naturally emitted behaviors (*operants*) or shaped them into more elaborate behaviors. Intervening variables and formal theory had no place in Skinner's *radical behaviorism*, which was more akin to Watson's than to Hull's thinking. A completely different, but still respectable, direction was to go inside the organism and work on *physiological psychology*. At Yale, for example, Neal Miller turned to studying the physiological underpinnings of learning, Frank Beach worked on neural and hormonal control of sexual and maternal behavior, and Karl Pribram did innovative work in neuroscience.

Yale's graduate students took various parts of Hull's legacy off in different directions; we will note three. First, some who were students in the 1940s continued the work on *verbal learning* that had been one of Hull's concerns, developing increasingly complex *mediation theories* (e.g., Osgood, 1953). The idea was to account for human language phenomena such as semantic generalization and transfer of training by positing chains of internal stimuli and responses that mediated between the observable ones. James J. Jenkins and colleagues (section 2.4) disconfirmed a key prediction in 1963, but optimism reigned during the 1950s. There was also an increased emphasis on methodology and measurement (e.g., it was shown that nonsense syllables varied in their meaningfulness, as indicated by the number of word associations they could produce; this became a factor to control in the design of experiments).

Second, some key students in the next generation, the early 1950s, gradually moved away from their Hullian roots in verbal learning to study memory, language, or visual imagery in the 1960s and beyond. George Mandler, for example, looking back on his Yale days (in Baars, 1986, p. 254), noted that "'Cognition' was a dirty word for us . . . because cognitive psychologists were seen as fuzzy, hand-waving, imprecise people who never really did anything that was testable." Nonetheless, Mandler became interested in the idea that memory was organized and made that one focus of his research at Harvard, Toronto, and then the new psychology department at the University of California, San Diego. Gordon H. Bower (a Neal Miller student) moved right from his Ph.D. at Yale to an influential career at Stanford, and Roger Shepard (a student of Hull's student, Carl Hovland) went to Bell Laboratories and Harvard University before settling at Stanford in 1968. All three were leaders in creating the cognitive psychology of the 1960s and beyond (see section 2.4).

Third, a set of researchers that overlapped somewhat with the erstwhile verbal learners revamped Hull's *mathematical modeling* strategy. The key contribution came from an unexpected person and place: William K. Estes, who had obtained his Ph.D. under Skinner at Minnesota and followed him to Indiana but was strongly influenced by Hull and Guthrie. His 1950 stimulus sampling theory was a less global but better motivated learning model that helped kick off two vigorous decades of work in mathematical psychology. Estes moved to Stanford in 1962, where he joined Richard C. Atkinson (who had been his student at Indiana), Patrick Suppes, Bower, and a new generation of students. Although Estes left in 1968, this critical mass of researchers at Stanford would play a major role in producing improved learning models. Other centers of activity included Indiana University (where Atkinson's student Richard M. Shiffrin made his career), Harvard and MIT in Cambridge (including, at various times, C. F. Mosteller, Robert R. Bush, R. Duncan Luce, and the ubiquitous George Miller), the University of Pennsylvania (where Eugene Galanter collaborated with Luce and Bush in the 1960s), and Bell Laboratories. It would take us too far from our main story to meaningfully describe the work produced by these mathematical psychologists in the 1950s. However, the same individuals played an important role in the transition to cognitive psychology in the 1960s, and we will meet some of them again in sections 2.1–2.4.

1.2.3 Alternatives during the era of behaviorism

In addition to diversity within the behaviorist camp, a variety of psychological endeavors thrived beyond the bounds of behaviorism. Many of these were situated in Europe, where behaviorism actually had little impact. We will briefly explore a few examples of nonbehaviorist research in the first half of the twentieth century that later contributed to the development of cognitive psychology.

We start in Britain with Sir Frederic Bartlett (1932), an experimental psychologist who studied the role of subjective construction in MEMORY. Memories, he claimed, are not simple recordings of experienced events, but are filled in by their subjects and embellished with details not present in the original context. For example, when asked to recall a Native American folktale, "The War of the Ghosts" from the Kwakiutl people, his subjects made changes in the plot of the story which tended to Westernize it. To explain this, Bartlett proposed that they employed their existing *schemata* to organize events in the story. As we will see, the notion of a schema as a structure for organizing information in memory has played a major role in subsequent cognitive psychology and in cognitive science generally. Bartlett also trained a number of influential British psychologists, including David Broadbent, who pioneered ATTENTION research using multi-channel listening techniques.

Beginning in the 1920s in Switzerland, Jean Piaget produced a huge and impressive body of work in the field he called *genetic epistemology*. Piaget's route into psychology was via his precocious studies of biology – he published at age 10 and received a doctorate at 21 – and his early and long-standing desire to integrate scientific and epistemological concerns. Pursuing postdoctoral studies in psychology and philosophy, he worked for a time in the Binet Laboratory in Paris and became intrigued by children's errors on standardized reasoning tests. He proceeded to devise ingenious methods for uncovering children's changing competencies, and over the years worked out an elaborate and unabashedly mentalistic theory that laid out stages of development and the internal processes responsible for children's movement through these stages. The theory was an edifice of twentieth-century thought that spawned hundreds of disciples, critics, and revisionists throughout the world, including many in North America (see Article 6, COGNITIVE AND LINGUISTIC DEVELOPMENT).

During the 1920s in the (then) Soviet Union yet another tradition with a cognitive or psychological flavor emerged from the group at the Institute of Psychology in Moscow. Lev Vygotsky developed a cultural-historical approach to psychology which guided his empirical work on children's cognitive and linguistic development; Alexander Luria maintained a like degree of theoretical breadth while focusing much of his empirical work on language disorders and functions of the frontal cortex. Particularly influential was their proposal that cognitive abilities emerge in interpersonal interactions (e.g., talking to others) before they assume a central role in private mental life (e.g., thinking to oneself in language). Today's inquiries into MEDIATED ACTION are rooted in this Soviet tradition.

An especially important counterpoint to America's behaviorism was the emergence in Germany and Austria of Gestalt psychology in 1912. Gestalt psychologists primarily studied PERCEPTION, especially perceptual and cognitive organization. They observed that the global properties of a whole object, such as its overall contour, are often more salient in perception than are component parts. A foundational study of the Gestalt movement was Max Wertheimer's (1912) examination of the so-called *Phi* phenomenon: the apparent motion when one light flashes on-off a split second after a nearby light has flashed on-off. Wertheimer argued against accounts of the Phi phenomenon according to which it was built out of separate recognition of the two lights flashing and offered an alternative account in terms of so-called field properties of the brain.

The notion that persons and animals often see or perceive things whole was a central conviction of Gestalt psychology (rather than a secondary theme, as it was for Wundt). In some cases the whole is spatial, as when perceiving the roundness of an object; in other cases it is temporal, as when an individual imagines goals and organizes behavior as a means to those goals. An example of perceiving temporal wholes is Wolfgang Köhler's research from 1913 to 1917 with chimpanzees in the Canary Islands. Köhler posed the problem of securing a banana that was out of reach, for example, and noted that chimpanzees solved the problem not by random trial and error, but rather by an overall reorganization of the parts of the situation that enabled an intelligent solution. This reorganization seemed to be discovered and implemented after a period of quiet planning. Hence, a chimpanzee would observe the situation and then go to a tree, tear off a branch, and use the branch to get the bananas. Otto Selz argued that such PROBLEM SOLVING required organizing the problem in a stepwise manner, with each step involving an operation on a representation of the problem (his work was influential in the development of Simon's thinking in AI). Köhler himself characterized such solutions as exhibiting insight (the *aha!* experience).

Returning to North America, the extent of behaviorism's reach varied with location and research area. In particular, the research area of sensation and perception felt little real impact from behaviorism – due in part to deep historical roots in a line of empirical inquiry going back to the mid-nineteenth century. Most researchers proceeded along lines laid down long before the emergence of behaviorism. Nevertheless, individual circumstances created niches outside the mainstream for a few novel approaches. First, the major proponents of Gestalt psychology relocated to such colleges as Swarthmore and Smith in the 1930s due to Hitler's rise to power in Europe. Some of their ideas and phenomena got picked up, but the theory as a whole gained few converts. Second, J. J. Gibson developed an ecological approach that Eleanor J. Gibson extended to perceptual development. Both were exposed to Gestalt psychology during their early careers at Smith College, and Eleanor Gibson did her Ph.D. work with Hull and Hovland at Yale; but they turned their backs on these approaches to pursue the idea that information and invariances in the environment are available for *direct perception*. The Gibsonian sphere of influence extended from their base at Cornell University to such places as Haskins Laboratory and the University of Connecticut, but in recent years it has broadened (see section 3.3).

Returning to mainstream research on sensation and perception, we said this research had deep historical roots. How deep? In one area, color vision, a major achievement in 1968 was Jameson and Hurvich's integration of two opposing theories first proposed by Hering and Helmholtz in the nineteenth century (see Article 19, PERCEPTION: COLOR). In another area, psychophysics, researchers seek to establish relationships between physical stimuli and subjects' subjective experiences of them. In the nineteenth century, Gustav Fechner (building on work by Ernst Weber) demonstrated that with respect to such variables as brightness or loudness, the perceived intensity of a stimulus is proportional to the logarithm of its physical intensity. Although perceived intensities are subjective, psychophysicists were able to develop methods for obtaining behavioral reports that lent themselves to systematic analysis. For example, a standard stimulus would be paired with a succession of comparison stimuli, with the subject reporting for each pair whether they appeared to be the same or different; from these reports Fechner calculated the *just noticeable difference (JND)* at various intensities and derived *Fechner's law* (the logarithmic scale noted above). With some refinements, methods of this kind continued to dominate psychophysics for a century. Then in 1956 at his Psychoacoustic Laboratory at Harvard, S. S. Stevens introduced the new, more direct method of magnitude estimation, in which subjects assigned a number to one stimulus at a time. Using this method he obtained *Stevens' law* (in which the scale is based on a power function). This type of scientific method and progressive refinement of laws was considered respectable, even to most behaviorists. It is not merely coincidental that several of the researchers most responsible for the early development of cognitive psychology, including George Miller, Ulric Neisser, and Donald A. Norman, received their Ph.D. training in psychoacoustics.

Next we briefly note three research areas that arose more recently and are more diverse than sensation and perception. First, developmental psychology in North America was pluralistic enough to serve as a seedbed and safe haven for the study of mental functioning at the same time that some developmental researchers pursued the implications of behaviorism and others limited themselves to descriptive studies. Piaget's influence on developmental psychology in North America was delayed until the 1960s, when researchers replaced his flexible *revised clinical method* with standardized procedures that confirmed his empirical findings. But even prior to this, Arnold Gesell's maturational approach and Heinz Werner's organismic-developmental psychology were available as alternatives to behaviorist studies of learning and conditioning in children.

Second, social psychology was even more pluralistic. One relatively cognitive line of researchers began with Kurt Lewin, a nonorthodox Gestalt psychologist who was among those relocating to North America in the 1930s. He was mentor to such major figures as Leon Festinger, who proposed his famous theory of cognitive dissonance in 1957. The claim was that subjects would modify their beliefs so as to reduce the inconsistency or dissonance between their beliefs and their behaviors. When they were unable to eliminate the dissonance, they would exhibit psychological discomfort. Just one of the many kinds of evidence which Festinger obtained involved the dissonance that smokers experienced as information began to appear in the 1950s that smoking causes lung cancer. He showed that heavier smokers who did not succeed in stopping were more reluctant to accept the evidence than were more moderate smokers who also could not stop.

Finally, the antecedents of today's clinical psychology should not be forgotten. There was considerably less specialization in the first half of the twentieth century than in the second half, and a surprising number of America's leading experimental psychologists devoted part of their careers to investigating such topics as personality and aptitude assessment, hypnosis, emotion, and psychodynamic theory (some had even been psychoanalyzed).

1.2.4 *Happenings at Harvard*

Although we have identified a number of alternatives to behaviorism, its influence, at least in North America, was powerful and widespread. Accordingly, a major transformation was required before psychology in the United States could become a contributor to cognitive science: it had to become cognitive again. Mediation theory had already prepared the way for rejecting the behaviorist proscription on appealing to mental events in explaining behavior; the cognitivists completed the job by also rejecting the stimulus–response framework for conceptualizing internal events. However, cognitivism retained other aspects of the behaviorist legacy: (1) its principle that behavior provided psychology's objective evidence, and (2) its methods for systematically gathering and analyzing that evidence (especially statistical significance testing and mathematical modeling). Cognitivists would posit mental events without apology, but only after predicting and confirming their effects on observable behaviors.

Many of the first stirrings of a cognitive psychology that would eventually overthrow behaviorist strictures originated with two Harvard psychologists, Jerome Bruner and George Miller. Harvard's administration split its psychologists between two departments in 1946 (and rejoined them in 1972–3): a new interdisciplinary Department of Social Relations and a reconstituted Department of Psychology with Edwin G. Boring as chair. In addition to the personality conflicts that occasioned the reorganization, the two departments provided homes to two very different theoretical orientations – Boring called them *sociotropy* and *biotropy*. Bruner and Miller found themselves, despite converging interests, in different departments. Bruner was in the Department of Social Relations. One of his first contributions was the development of the New Look movement in the psychology of PERCEPTION, which (a) emphasized the contribution of internal mental states (partly determined by social factors) of the perceiver to what is perceived and (b) denied that the external stimulus is *the* determining factor in perception. In the psychophysics that dominated the study of perception at the time, variability in perceptual judgments was regarded as an impurity, but Bruner brought variability to center stage.

In 1947 Bruner and Cecile Goodman, a Harvard undergraduate, performed a study showing that children's judgments of the sizes of coins varied with their value: the size of lower-valued coins was underestimated, while that of higher-valued coins was overestimated. This contradicted the general principle of the psychophysical Law of Central Tendency, according to which judgments of smaller and larger items should err in the direction of the mean or central tendency in a series. Further, the overestimates of higher-valued coins were even larger for poorer children, revealing a social effect on perception.

Most of the major studies that defined the New Look were done collaboratively by Bruner and Leo Postman. Together they began to study the ability of subjects to read words flashed quickly through a tachistoscope. They discovered that the time required to read a word varied with a number of factors, including whether the word was closely associated with values that were strongly held by the subject. What was surprising about this and related findings was that in some way the semantic significance of a word could affect processes prior to the actual recognition of the word itself. Equally surprising were the results of their 1949 study of tachistoscopic perception of playing cards, some of which were anomalous in that the color was reversed (e.g., a red 10 of clubs). For each card, very brief exposure times were lengthened until it was recognized. With the anomalous cards, subjects would initially respond with a suit appropriate to the color displayed (e.g., a red 10 of clubs would be seen as a 10 of diamonds). Only at very long durations were they able to recognize the anomalous combination of suit and color. After experience with some of the anomalous cards, though, subjects learned to recognize them as rapidly as the normal cards. This set of experiments revealed both the role of expectations in perception and the possibility of learning to see new things or old things in new ways. (At the time of the Bruner–Postman experiments, Thomas Kuhn was a Harvard Fellow. In his 1962 book, *The Structure of Scientific Revolutions*, he employed the New Look results as evidence against the claim that scientists are purely objective reporters; rather, they may fail to see certain phenomena until a new paradigm changes their expectations.)

Bruner soon turned from showing the role of thinking in perception to a more direct exploration of thinking, which culminated in his 1956 book with Jacqueline Goodnow and George Austin, *A Study of Thinking*. Bruner and his colleagues regarded the learning and use of categories as central to thought. To investigate category learning, they built on a procedure from Lev Vygotsky using arrays of cards with geometrical patterns (e.g., two black circles surrounded by a single border). The investigator mentally chose a rule defining a category (e.g., all cards with two circles), and subjects tried to discover the rule by picking one card at a time for the investigator to identify as an exemplar or nonexemplar of the category. Each subject's sequence of selections could then be examined to determine what strategies were being used. A common, generally successful strategy was for the subject to find one positive instance of the category, and then to systematically pick other cards that differed in one attribute: if the new card was also a member of the category, then the dissimilar attribute should not be relevant to the category assignment. This program of exploring thought through the window of category reasoning was expanded in a variety of important and interesting directions in subsequent years. For example, in the 1970s Peter Wason and Philip Johnson-Laird conducted studies which were interpreted as showing that subjects manifested a *confirmation bias* in category-reasoning tasks, continuing to examine instances that would support a hypothesized rule rather than seeking out cases that might falsify it. Also in the 1970s Eleanor Rosch, one of Bruner's graduate students, began a program of study of natural categories which revealed that generally they lacked defining rules such as Bruner and his colleagues employed in their studies of contrived categories (see Article 8, CONCEPTUAL ORGANIZATION, and Article 25, WORD MEANING). The foundation for these and many other developments, however, was the seminal work reported in *A Study of Thinking*.

While Bruner carried out his work on perception and categorization in Harvard's Department of Social Relations, George Miller was at work in the Department of Psychology. That department was dominated by two personalities: B. F. Skinner, who returned to Harvard from Indiana in 1948, and S. S. Stevens. During World War II Miller worked in Stevens's laboratory on optimal signals for spot jamming of speech, which was classified military research. At his Ph.D. oral, Miller had to present his results very discreetly; accordingly, he spoke not of jamming, but rather of the effects of noise on intelligibility of speech. One of his observations was that certain messages were harder to jam (or easier to understand in noisy environments) than others, a finding that did not seem explicable in terms of the physical acoustical data alone but was soon to receive a novel interpretation. After receiving his Ph.D. in 1946, Miller remained at Harvard until 1968 (except for the early 1950s, when he spent a year in Princeton and then four years at MIT). Miller's interest in mathematical analysis led him to actively follow developments in statistics and information theory, especially Shannon's (1948) paper on information theory.

Claude Shannon, whom we have already encountered as a designer of switching circuits that implemented Boolean operations, was a mathematician at Bell Telephone Laboratories. This organization had a natural interest in understanding the laws governing transmission of information and especially in determining the maximum quantity of information that could be transmitted over an acoustic channel. Shannon's theorizing about information started from the observation that the quantity of information transmitted over a channel depends on variation in a signal. In the simplest case, the variation would involve just two equally likely alternatives: *on* or *off*. Then the basic unit of information, a *bit* (binary unit), could be defined as the amount of information transmitted by selecting one of these alternatives rather than the other (a binary decision). A single bit can distinguish between just two alternatives (1 corresponds to *on*, 0 to *off*). Adding a second bit (an additional *on/off* signal) doubles the number of alternatives that can be distinguished to four (00, 01, 10, 11). A third bit doubles the information again – a three-bit sequence distinguishes among eight alternatives. (Wendell Garner, who played a major role in applying Shannon's ideas to psychology, pointed out in 1988 that this approach defines information in terms of all the possible events that could have occurred, not just the actual event. For example, the informativeness of the event 10 – *on* then *off* – depends on the fact that it excludes exactly three other events.)

This general approach can be extended to more complex situations in which there are more than two alternatives or alternatives have unequal probabilities – for example, any message in English – and can be used to measure the amount of redundancy in such messages. Shannon (1948, 1951) presented a text one letter at a time to subjects whose task was to predict the next letter. There were 26 alternatives at each point, and they had unequal probabilities due in part to context. For example, *x* has a low probability overall, but is highly probable following *mailbo*. Shannon defined redundancy as the reciprocal of the average number of guesses needed to generate the correct letter. Averaging across the entire text, subjects required an average of two guesses per letter, yielding a redundancy estimate of about 50 percent for printed English. Shannon's information theory provided the key to interpreting Miller's dissertation result that messages differed in how easily they could be understood in noisy environments. Miller and Selfridge (1950) found further application for information theory in a list learning experiment: the closer the word lists came to resembling English sentences (i.e., the greater their redundancy), the more words a subject could remember.

In one of the most influential papers of this period, Miller (1956) addressed more extensively the question of the cognitive structure of MEMORY. The study of human learning and memory had long moved along the path laid down by Hermann Ebbinghaus (1885/1913), who served as his own subject in a prolonged series of experiments in order to bring higher mental processes under experimental control and quantitative analysis. In his attempts to eliminate extraneous influences, Ebbinghaus arrived at the idea of using pronounceable nonsense syllables such as DAX and PAF as his stimuli rather than words. He studied lists of these nonsense syllables daily, and then tested himself to determine rates of learning and forgetting. Ebbinghaus uncovered important functional relations (e.g., repetition yields better retention, especially if distributed across several days; the amount retained is a logarithmic function of time), but the down side was his neglect of the cognitive structures and processes that meaningful stimuli so readily engage. Frederick Bartlett's (1932) previously described idea that schemata help organize memory offered a corrective to the limitations of the Ebbinghaus tradition, but Bartlett's early impact was felt primarily in England. Verbal learning in North America (as discussed in section 1.2.2) pursued updated variations on the Ebbinghaus tradition by asking, for example, which particular model of stimulus–response conditioning might best account for the accumulated data on paired-associate learning. Retention was an indicator of learning, not a clue to the nature of the memory system within.

Miller made an exploratory claim about the underlying memory system by pointing in his 1956 paper to an interesting limitation. Over a short period of time, humans can retain only about seven items in memory ("the magical number seven, plus or minus two"). This limit could be overcome if the items formed coherent units, or chunks (as do the letters of a familiar word or acronym). Thus, the sequence of letters I, B, M, C, I, A, B, B, C, U, S, A exceeds the limit, and so is very difficult to remember. When it is chunked as IBM, CIA, BBC, USA, it falls well within the limit and is easy to remember. The limit reemerges, however, in that humans can retain only about seven chunks unless those can themselves be rechunked. For Miller, the limitation to seven items was not an isolated, odd fact. He began his paper: "My problem is that I have been persecuted by an integer. For seven years this number has followed me around, has intruded in my most private data, and has assaulted me from the pages of our most public journals" (1956, p. 81). For a variety of activities, such as distinguishing phonemes from one another, making absolute distinctions amongst items, as well as remembering distinct items, when the number of items reached seven, significant changes arose. Miller concluded: "There seems to be some limitation built into us either by learning or by design of our nervous system, a limit that keeps our channel capacities in this general range" (p. 86). Miller's essay became a classic of cognitive psychology, in part because it suggested that there was structure to the internal processing system, whose character could partly be discovered via **BEHAVIORAL EXPERIMENTATION.**

In the research of Bruner and Miller, an approach to psychology was stirring which retained the emphasis on behavioral data but rejected behaviorism's suspicion of mental machinery. The enterprise would not be named *cognitive psychology* until 1967, but its *modus operandi* – seeking structure in the mind which could explain features of behavior – was already apparent.

1.3 *The brain develops*

After cognitive science left its conceptual womb, as it were, and began to mature, it went through a long period in which contributions from neuroscience were either ignored or actively dismissed as irrelevant to the pursuits of cognitive science. But while it was in the womb during the 1940s and 1950s, advances in understanding the brain contributed to researchers' thinking about how concepts such as information and computation might provide a basis for understanding mental processes. We have already observed links between psychology and work on computation in the cybernetics movement. The idea that psychology could benefit from collaborations with neuroscience as well was articulated by psychologist Donald Hebb of McGill University in the preface to his classic 1949 book, *The Organization of Behavior*:

> There is a considerable overlap between the problems of psychology and those of neurophysiology, hence the possibility (or necessity) of reciprocal assistance. . . . Psychologist and neurophysiologist thus chart the same bay – working perhaps from opposite shores, sometimes overlapping and duplicating one another, but using some of the same fixed points and continually with the opportunity of contributing to each other's results. (pp. xii and xiv)

The spirit of Hebb's book was clearly evident in the September 1948 conference on "Cerebral Mechanisms in Behavior" sponsored by the Hixon Fund. Speakers included neurophysiologists Warren McCulloch and Rafael Lorente de Nó; biologically oriented psychologists Ward Halstead, Heinrich Klüver, and Karl Lashley; Gestalt psychologist Wolfgang Köhler; and computer scientist John von Neumann. The papers covered a wide range of topics, including von Neumann's analysis of the similarity of computers to the brain, Köhler's study of evoked potentials during pattern perception, Klüver's comparison of functional contributions of the occipital and temporal lobes, and Halstead's attempt to relate intelligence to the brain. As the discussion recorded in Jeffress (1951) makes evident, the divides between psychologists and neuroscientists were not large, and discussion flowed easily from psychological phenomena to neural mechanisms.

While disciplines such as psychology and linguistics were in a position to contribute to the birth of cognitive science only after undergoing internal revolution, and artificial intelligence had first to be created, neuroscience had a much longer and continuous history. The idea that the brain was not merely the organ of mental processes but might be decomposed into component systems which perform different specific functions in mental life was a product of the nineteenth century. The challenge to neuroscience, then and now, is to parse the brain into its functional components and (a more difficult task) to figure out how they work together as a system. The implication for cognitive science is that information about the distinct functions performed in the brain could be used to corroborate or guide the development of psychological models of cognitive activities. Exploring this functional decomposition and localization depended in part on the development of appropriate tools. What follows is a very selective review of research from the nineteenth century through 1960 that began to determine the structure of the brain and the relation of its components to mental life.

1.3.1 *Neural architecture*

Before scientists could make claims about the functional organization of the brain, they needed to learn something about its general architecture. At the end of the nineteenth century major advances were made at both the micro and the macro level in understanding the brain. At the micro level the crucial breakthrough was the discovery that nerve tissue is made up of discrete cells – neurons – and that there are tiny gaps between the axons that carry impulses away from one neuron and the dendrites of other neurons that pick up those impulses. In the 1880s Camilio Golgi introduced silver nitrate to stain brain slices for microscopic examination. Silver nitrate had the unusual and useful feature of staining only certain cells in the specimen, thereby making it possible to see individual cells, with their associated axons and dendrites, clearly. Using this stain, Santiago Ramón y Cajal argued that the nervous system was comprised of distinct cells (a view that Golgi, however, never accepted). Sir Charles Scott Sherrington then characterized the points of communication at the gaps between neurons as *synapses* and proposed that this communication was ultimately chemical in nature.

While processes at the micro level of the neuronal substrate would figure prominently in understanding cognitive processes such as learning (which is widely thought to involve changes at synapses that alter the ability of one neuron to excite or inhibit another) and became the inspiration for computational modeling using neural networks (see section 1.3.5), another kind of advance involved linking different macro-level brain areas with specific cognitive functions. This required overcoming the view widely shared in the eighteenth century that the brain, especially the cerebral cortex, operated holistically, without any localized differentiation of function.

The major credit for promoting the idea that the macrostructure of the brain was divided into distinct functional areas is due to Franz Joseph Gall. Working in the early nineteenth century, he proposed that protrusions or indentations in the skull indicated the size of underlying brain areas. He further thought that it was the size of brain areas that determined how great a contribution they made to behavior. Accordingly, he proposed that by correlating protrusions and indentations in individuals' skulls with their excesses and deficiencies in particular mental and character traits, he could determine which brain areas were responsible for each mental or character trait. *Phrenology*, the name given to Gall's views by his one-time collaborator Johann Spurzheim, has been much derided as quackery. Nonetheless, Gall's fundamental claim that differentiation of structure corresponds to differentiation of function in the cortex came to be widely accepted, so much so that those espousing localization of function in the latter part of the nineteenth century were often referred to as neophrenologists.

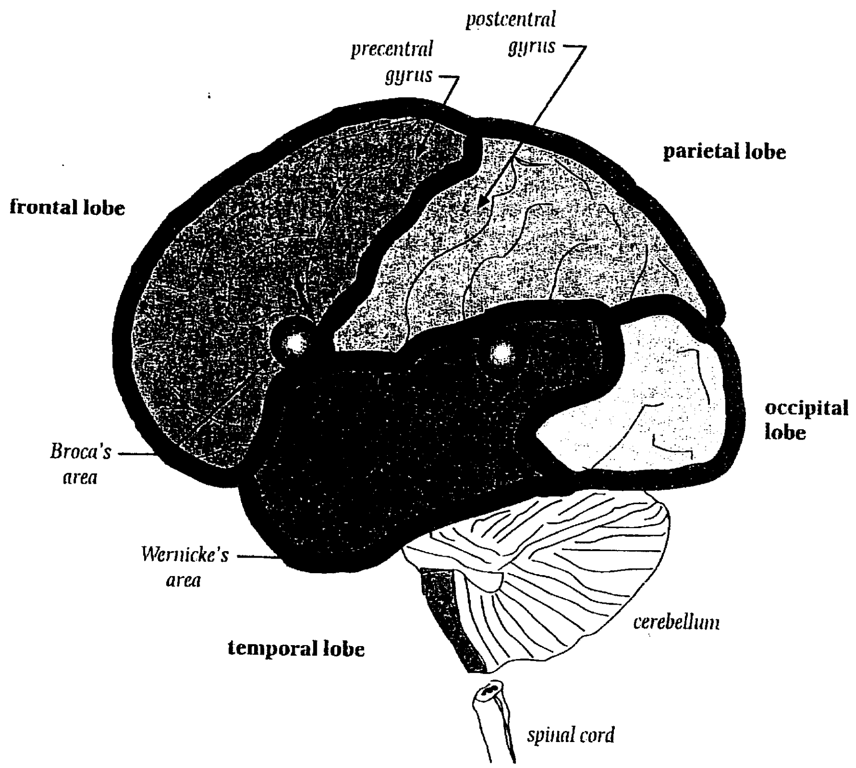

Figure I.3 Major features of the brain's architecture. Shown are the cerebellum and the four lobes of the cortex: frontal, parietal, temporal, and occipital. Also indicated are the approximate locations of Broca's and Wernicke's areas and the precentral gyrus (primary motor cortex) and the postcentral gyrus (primary sensory cortex). Figure by Robert S. Stufflebeam.

One problem that researchers faced in attempting to localize mental functions in the brain was the lack of any standardized way of designating parts of the brain. The folding of the cortex creates *gyri* (hills) and *sulci* (valleys); anatomists have named some of them and used the most prominent sulci to divide the brain into different lobes, as shown in figure I.3. But each lobe itself contains a number of anatomically distinct regions. Using such criteria as responses to various stains and the distribution of cells between cortical layers, a number of researchers at the end of the nineteenth century produced more detailed atlases of the brain. Of these, that by Korbinian Brodmann (1909) became the most widely adopted, and his numbering of brain regions is still widely employed today (see Article 4, BRAIN MAPPING).

A guiding principle for Brodmann and others who tried to map the brain was that different brain regions would perform different functions. The principle of localization of function has had vocal opponents. In the 1820s, Marie-Jean-Pierre Flourens voiced objections to Gall, while a century later neuropsychologist Karl Lashley was the most prominent critic of Brodmann. Arguing that higher cognitive processes (ones involved in memory and learning) were not localized but distributed, Lashley (1929) introduced two alternative principles: equipotentiality and mass action. *Equipotentiality* refers to the ability of brain regions to take on different functions as needed (e.g., if the region that previously performed a function were damaged), while *mass action* refers to the idea that the ability to perform higher functions relates to the total available cortex, not to any one part of it. In 1950 Lashley presented his arguments against localization of specific memory traces in the very often cited paper "In search of the engram," in which he recounted the repeated failures to localize such major functions as habitual behavior.

Despite the doubts of Flourens, Lashley, and others, most researchers have assumed that – at some level of detail in the analysis of function – functions are localized in the brain. To obtain evidence for particular localizations, researchers have had to develop a number of research techniques. We briefly review a number of them and some of the more prominent results obtained by using them.

1.3.2 *Deficit studies*

One of the earliest and most fruitful sources of information about the function of brain areas is the study of deficits that ensue when a neurostructure is damaged (see Article 29, DEFICITS AND PATHOLOGIES). The path from damage to conclusion is difficult to traverse, however: anatomical and functional brain areas vary across individuals, and it is also tricky to determine precisely what contribution the damaged part makes to normal function. Some of these challenges are well illustrated in the *locus classicus* of deficit studies, Paul Broca's (1861/1960) famous case of Monsieur Leborgne (more commonly known as "Tan" for the one syllable he could utter). When Tan was brought to Broca, he suffered not only loss of articulate speech, but also epilepsy and right hemiplegia. Tan died six days after Broca first saw him. Broca performed an autopsy, which revealed massive damage to the left frontal lobe. The central region of this lesion came to be known as *Broca's area* (see Article 13, LANGUAGE EVOLUTION AND NEUROMECHANISMS).

Broca's research led to a much more positive response to claims that mental functions are localized in the cortex. But not everyone agreed that mental functions should be identified with the regions in which damage could lead to loss of function. The variability in relations between structures in the brain and mental functions was emphasized by Charles-Edouard Brown-Séquard who, in response to Broca, presented atypical cases in which aphasia developed after damage to the right frontal cortex, cases in which lesions outside Broca's area affected speech, and cases in which damage to Broca's area did not result in speech deficits. Even some of those who accepted the claim that damage to specific areas would result in specific functional deficits rejected the claim that the function itself was performed in that area. This view is exemplified by Carl Wernicke (1874), who, in the decade after Broca, presented cases in which damage to an area in the temporal lobe that came to be known as *Wernicke's area* resulted in a loss of comprehension of speech while leaving the ability to speak intact. Wernicke, however, operating out of an associationist framework, viewed the site of damage as a locus of association between simple ideas, not the locus where comprehension was achieved.

Wernicke's associationism was largely overlooked. The standard view, one espoused especially by Norman Geschwind (1974), was that Wernicke's area is the locus of speech comprehension and Broca's area the locus of speech production. More recently, this decomposition of function into comprehension and production was challenged by researchers influenced by Chomsky. In a Chomskian LINGUISTIC THEORY there are separate components for phonology, syntax, and semantics. The theory is intended to be neutral with respect to comprehension and production. In 1980 Bradley, Garrett, and Zurif produced evidence that lesions in Broca's area produce deficits of syntactic processing that, although more obvious in production, can be observed in comprehension as well. This gave rise to a new standard view, which localized semantics in Wernicke's area and syntax in Broca's area.

The deficits discussed so far stemmed from natural lesions. But in the first half of the twentieth century surgeons sometimes induced lesions in the brain to alleviate effects of diseases such as epilepsy. Some of these patients subsequently experienced unanticipated effects, which provided suggestive evidence of the functional significance of the structure that had been lesioned. One prominent example of such surgery in the 1950s involved a patient who has become known in the literature by his initials, H.M. (Milner, 1965). In 1953, while in his twenties, H.M. had his hippocampus removed. This resulted in *anterograde amnesia*, loss of memory for events that happened after his surgery. H.M. could learn new skills, but could not learn new facts or events (e.g., he could not recall ever meeting physicians who saw him on a daily basis). This indicates that the hippocampus plays a crucial role in the storage of new information in memory; however, since H.M. retained previous memories, it is presumably not the site of the memories themselves. Also, the specificity of the deficit (H.M.'s intelligence was unaffected) constituted strong evidence against Lashley's doctrine of mass action.

1.3.3 Stimulation studies

From the fact that a lesion is accompanied by the loss of a function, one cannot conclude that the site of the lesion is itself the locus of the function; it may merely be the site of some necessary but relatively minor component of the function. One way of ameliorating this limitation of deficit studies is to augment them with information obtained by stimulating a specific part of the brain and observing the response. The classical example of this approach occurred just a few years after Broca's observations. In 1870 Gustav Fritsch, an anatomist, and Eduard Hitzig, a psychiatrist, collaborated on a study in which they applied low levels of electrical stimulation to different cortical areas of a dog. (Studies of electrical stimulation had been common earlier, but generally suffered from employing too strong an electrical current.) They found that stimulating specific sites on the cortex resulted in muscular responses on the opposite side of the body in the forepaw, hindpaw, face, and neck. They combined this stimulation study with a lesion study in which they removed the portion of the cortex that activated the forepaw; the result was impairment (although not total loss) of movement in the forepaw. Fritsch's and Hitzig's stimulation studies were followed up by David Ferrier, who performed comparative studies on a wide range of species and developed techniques for eliciting relatively fine movements such as the twitch of an eye.

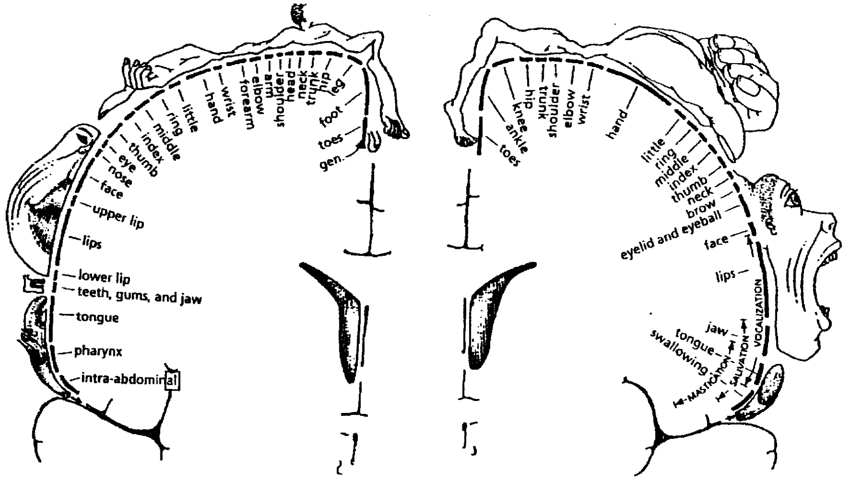

Figure I.4 Representations of the primary sensory cortex (left) and primary motor cortex (right), using homunculi to show the relative amounts of brain tissue devoted to different portions of the body. Redrawn from Penfield and Rasmussen (1950).

Stimulation studies gained practical applications as neurosurgery became more common in the twentieth century. The Canadian neurosurgeon Wilder Penfield, for example, developed a procedure of removing the portion of the brain in which epileptic attacks seemed to originate in severe epileptics. To reduce the risk of causing cognitive or behavioral deficits as a result of these surgeries, Penfield would insert tiny electrodes and induce a small current into the site of surgery to see what responses the patients would make. Employing patients' verbal reports of sensation as well as their own observations of motor responses, Penfield and Rasmussen (1950) were able to develop far more detailed maps than had been possible previously for the primary motor cortex and primary sensory cortex (located on the precentral gyrus and postcentral gyrus respectively). Figure I.4 illustrates the most striking feature of these maps: different parts of the body are represented by disproportionate amounts of brain area. The hand and face are represented in great detail, while representations of the trunk and neck are localized in much smaller areas. In other research Penfield stimulated areas of the temporal lobe; this often triggered highly detailed memories – memories more detailed than subjects would recall voluntarily. Although critics often emphasized the variability of these results and the fact that reports of supposed memories seemed to be highly related to current topics of conversation (see Valenstein, 1973), Penfield's observations excited many researchers with the prospect of localizing specific cognitive states in the brain. The stimulation technique also produced controversial claims of centers for hunger, aggression, pleasure, and so forth.

1.3.4 *Single neuron electrophysiology*

One of the major advances around 1950 was the development of SINGLE NEURON ELECTROPHYSIOLOGY – the technique of using microelectrodes to record the activity of individual nerve cells. Some of this research took place with simple organisms, as in Ratliff and Hartline's (1959) research on the retina of the horseshoe crab. Stimulating photoreceptor cells with bars of light and measuring activity in the ganglion nerve cells to which they were connected, the researchers found that the ganglion cells were more responsive to the edge between light and dark than to the center of the light bar. They explained this by positing that the *excitatory* projection from an aligned receptor was only one determinant of a ganglion cell's activity. There were also *inhibitory* connections from the neighboring receptors to the same ganglion cell (*lateral inhibition*), and the ganglion cell's response depended on the difference between the excitatory and inhibitory inputs. Ganglion cells connected to receptors near the edge of the light bar would receive as much excitation as those connected to receptors at the center of the light bar, but less lateral inhibition. Making the maximum response, they functioned as *edge detectors*.

Working with cats rather than crabs during the same period, but also using light bars as stimuli, David Hubel and Torsten Wiesel (1962) recorded activity further along the neural pathways involved in vision. They found cells in the visual cortex that were maximally responsive to light bars of particular width and orientation in specific locations in the visual field. They also identified cells that responded to light bars of particular width and orientation but were much less restrictive as to location. Hubel and Wiesel suggested that there might be a hierarchy of processing cells, such that cells detecting lines at various specific locations would feed into higher-level cells that detected lines at any of those locations. One of the best-known exemplars of this line of research was the demonstration by Jerome Lettvin, Humberto Maturana, Warren S. McCulloch, and Walter Pitts (1959) of cells in the frog's brain that responded to small dark roundish shapes in motion, and hence seemed to serve as *bug detectors*. Findings of this kind led to speculative discussions of the extent of localization and specialization of encoding in the human brain, with so-called *grandmother detectors* at one end of the spectrum and Lashley's mass action principle at the other.

1.3.5 *Computational modeling: neural networks*

A major new development in the 1940s and 1950s was the emergence of computational analyses of neural systems and the beginnings of brainlike computational modeling (an approach which, via the mediation of Donald Hebb, took over the term *connectionism* from earlier, associationist approaches to conceptualizing the brain such as Wernicke's). A key figure in this development was Warren McCulloch, a neurophysiologist who began his career at the University of Chicago. He collaborated with Walter Pitts, then an 18-year-old logician, in a widely cited 1943 paper that analyzed networks of neuron-like units. McCulloch and Pitts showed that these networks could evaluate any compound logical function and claimed that, if supplemented with a tape and means for altering symbols on the tape, they were equivalent in computing power to a universal Turing machine. The units of the network were intended as simplified model neurons and have been referred to ever since as *McCulloch–Pitts neurons*. Each unit is a binary device (i.e., it can be in one of two states: *on* or *off*) that receives excitatory and inhibitory inputs from other units or from outside the network. The state of a network of these units emerges over a number of cycles. On a given cycle, if a unit receives any inhibitory input, it is blocked from firing. If it receives no inhibitory input, it fires if the sum of equally weighted excitatory inputs exceeds a specified threshold. A unit with this design is appropriate not only as a model of a simplified neuron but also as a model of an electrical relay – a basic component of a computer – and hence McCulloch–Pitts neurons helped draw the connection between the brain and computers that was emphasized by others, including John von Neumann and Marvin Minsky. McCulloch and Pitts also made a link to logic: the neurons could be associated with propositions, and because of the binary nature of these units, their activation states could be associated with truth values.

As attractive as some theorists found the comparison of the brain to a computer at the architectural level, many others moved beyond the logic-gate level of focus and began trying to analyze how nervous systems carried out more complex psychological tasks, such as those of perception. These ambitious researchers included Pitts and McCulloch themselves, who, in a 1947 paper, tackled two knotty problems: how someone can recognize an object as the same when it appears in different parts of the visual field and how the superior colliculus is able to transform spatial maps of sensory inputs into motor maps that direct such activities as eye movements. Here they abandoned the earlier paper's focus on propositional logic in favor of spatial representations and analog computations. A further departure from the earlier paper is an emphasis on networks that rely on statistical order and operate appropriately despite small perturbations. Moreover, as part of their evidence for specific computational models, they compared diagrams of these computational systems with diagrams of specific neural structures.

The focus on perception continued in the central parts of Donald Hebb's 1949 book, *The Organization of Behavior*. The subtitle, "Stimulus and response – and what occurs in the brain in the interval between them," points to one of the main emphases of Hebb's analysis: the development of internal structures that mediate stimulus and response. Hebb sought to overcome the opposition between the more localizationist switchboard theories emphasizing sensory–motor connections and the anti-localization approaches of the Gestalt theorists and his own mentor, Lashley. The key to his alternative was the notion of neuronal cell assemblies, which consisted of interconnected, and hence self-reinforcing, sets of neurons which represent and transform information in the brain.

> Any frequently repeated, particular stimulation will lead to the slow development of a "cell-assembly," a diffuse structure comprising cells in the cortex and diencephalon (and also, perhaps, in the basal ganglia of the cerebrum), capable of acting briefly as a closed system, delivering facilitation to other such systems and usually having a specific motor facilitation. A series of such events constitutes a "phase sequence" – the thought process. Each assembly action may be aroused by a preceding assembly, by a sensory event, or – normally – by both. The central facilitation from one of these activities on the next is the prototype of "attention." (1949, p. xix)

Hebb proposed that these subassemblies were created by an interaction between cells whereby every time one cell figured in the firing of another, their connection was strengthened: "When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased" (p. 62).

One of the first attempts to model neural processes on a digital computer was Rochester, Holland, Haibt, and Duda's (1956) study of Hebb's proposal for cell assemblies, carried out at IBM. They discovered the need for several additions and modifications to Hebb's proposal, including (a) inhibitory connections within the net and (b) a mechanism for normalizing connection weights so that they would not grow without bound. Perhaps most importantly, they modified the learning rule so that connection weights would be reduced when one unit was active but the other was not. The modified Hebb rule has become one of the standard *learning rules* employed in subsequent connectionist modeling (see Article 16, MACHINE LEARNING).

The culmination of this early work on modeling the brain was psychologist Frank Rosenblatt's work on perceptrons at Cornell (Rosenblatt, 1962). Figure I.5 shows a simplified perceptron (omitting backward and lateral connections). It had an input device or retina consisting of a number of binary units. Each of these had weighted connections to a number of other units (associator units), which would become active whenever the combined activation from the input units exceeded a threshold. These associator units were connected in turn to response units. The perceptron's task was to activate the appropriate response units when a pattern was activated on the input units. The perceptron's ability to do this depended on its having appropriate weights on its connections – those that would produce the intended classification of the input patterns.

One of Rosenblatt's major contributions was to develop a learning procedure for adjusting the strength of the connections from the associator units to the response units. The procedure began by supplying a pattern to the input units and allowing the network to generate a response. The activity of each response unit was then compared with the desired response. If the unit was on when it should have been off, then each connection from an active associator unit was weakened; if the unit was off when it should have been on, each connection from an active associator unit was strengthened. Rosenblatt proved the Perceptron Convergence Theorem, which established that this procedure would succeed in finding connection weights that would permit the network to produce correct responses unless no such set of connection weights existed. Rosenblatt discovered that there were some sets of input–output pairings which no set of connection weights could generate. He explored a number of variations, including networks with feedback from response units to associator units and networks with multiple layers of associator units. He also explored procedures through which the network would self-organize as a result of inhibitory connections between units, hence learning response patterns without a trainer. These projects were never completed; Rosenblatt died in his early forties in 1971.

Many of Rosenblatt's ideas have been taken up by more recent neural network researchers (see Article 38, CONNECTIONISM, ARTIFICIAL LIFE, AND DYNAMICAL SYSTEMS), but one thing that was distinctive about Rosenblatt's work was that he built actual physical devices, rather than limiting himself to mathematical analyses of hypothetical devices or simulations on digital computers. One of these devices was the Mark I Perceptron, which had a 20 by 20 input grid of photoelectric cells, 512 associator units, and 8 response units. The variable connections between associator and response units were implemented by motor-driven potentiometers. Some of the experiments conducted on the Mark I included studies with noisy target patterns or with damaged networks (accomplished by removing actual wire connections). In a subsequent model, named "Tobermory" after his own cat, Rosenblatt ambitiously attempted to model the visual system of a cat.

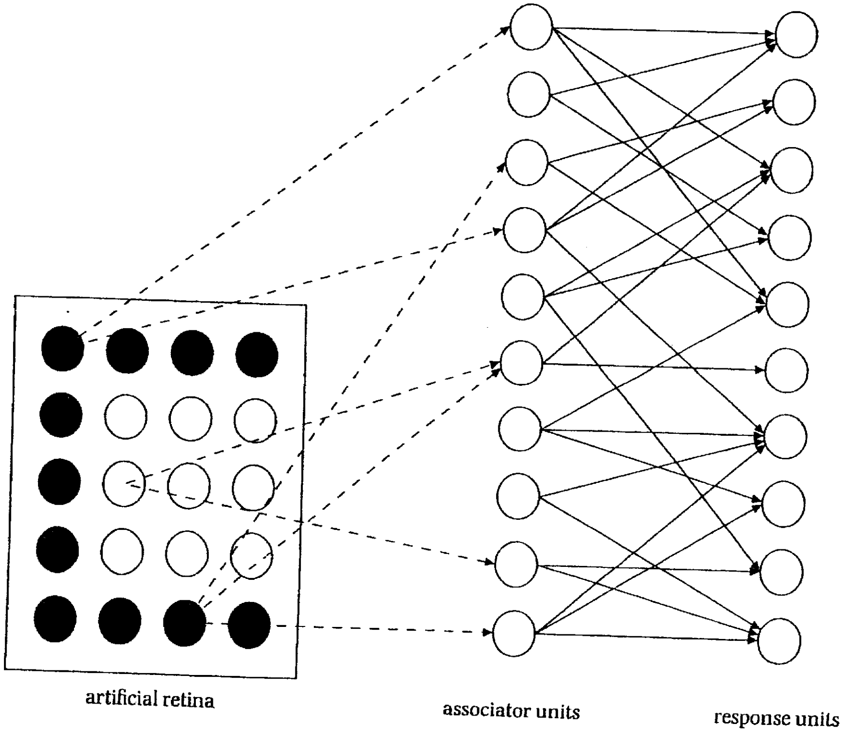

Figure I.5 A simplified perceptron. Each unit in the artificial retina is connected to a randomly selected set of associator units. Each associator unit is in turn connected to a randomly selected set of response units. Rosenblatt (1962) fixed the weights of the first set of connections (dashed lines) but developed learning procedures for the second set of connections (solid lines).

1.4 *Viability: the transformation of linguistics*

Like psychology, linguistics had to undergo a transformation in order to make its contribution to cognitive science. The central figure in this transformation was Noam Chomsky; but before turning to his innovations, we should consider briefly the nature of linguistics prior to Chomsky. In a landmark achievement that began in the late eighteenth century, *historical linguists* reconstructed the long extinct Proto-Indo-European language and traced its divergence into contemporary languages of Europe, Iran, and northern India. In the late nineteenth century, largely inspired by the Swiss linguist Ferdinand de Saussure, the focus shifted: *structural linguists* began to describe the current structure of particular languages rather than their history. This involved analyzing a corpus to identify units at several levels. For example, the word *boats* is composed of two morphemes, one composed of three phonemes, /bot/, and one composed of one phoneme, /s/. European structuralists tended to take an explanatory stance and to emphasize the role of language in conveying meaning. North American structuralists, by contrast, became descriptivists with a vengeance. Franz Boas (a contemporary of Saussure, but educated in anthropology), his student Edward Sapir (who combined Boas's descriptivism with broader, almost European concerns), and Leonard Bloomfield (the previously mentioned behaviorist linguist of the 1920s through 1940s) put considerable emphasis on developing appropriate methods for identifying the phonemes and morphemes of any language. With Native American languages rapidly losing their speakers, they felt some urgency in applying these methods to the analysis of as many indigenous languages as possible.

Bloomfield's distinctive contribution was his positivist stance, reflected in his affinity for psychological behaviorism and his insistence on a rigorously empiricist linguistic methodology: "Accept everything a native speaker says in his language and nothing he says about it." Bloomfield's influence was enhanced by his energy for organization: he played a key role (with Boas and Sapir) in creating the Linguistic Society of America in 1924 and sought to focus it on the "science of language" (versus such pursuits as literary analysis) in its journal, *Language*, and in the courses offered in its summer Linguistic Institute. These are still the premier institutions in American linguistics, but they played an especially critical role in defining the field during an era in which most linguists were employed in language or anthropology departments. By the 1940s and 1950s, Bloomfield's students and followers – the *post-Bloomfieldians* – dominated linguistics. They prided themselves (overly optimistically) on their near completion of a system of *discovery procedures* – an objective sequence of operations sufficient to arrive at a phonological and morphological analysis of any language. New technologies even created the dream that these procedures might be mechanized. The sound spectrograph was invented at Bell Laboratories and made public in 1945; its visual display of speech frequencies across time fostered a surge of new knowledge concerning acoustic phonetics that had implications for linguistics. Digital computers gave rise to machine translation projects in the early 1950s (including Victor H. Yngve's group at MIT) and to early parsers in the late 1950s (including one programmed under the direction of Zellig Harris; see Joshi & Hopely, 1997, for an account).

Meanwhile, new ideas about language were flowing from two adjacent fields. First, the counterparts of the post-Bloomfieldians in psychology – Hull's successors – were pursuing mediation theories of language (section 1.2.2). Second, information theory was suggesting ways to approach speech statistically (section 1.2.4). In retrospect, both these approaches are notable for their transitional character. Mediation theory was constructed in an S–R framework, but drew ever closer to mentalism. Information theory contributed the idea of information, but as applied in cognitive science, its statistical approach to information was largely superseded.

Hull died in 1952, and Bloomfield in 1949. Their successors, enjoying optimistic times in their respective disciplines of psychology and linguistics, were ready to rediscover the advantages of interdisciplinary cooperation. At an eight-week summer seminar sponsored by the Social Science Research Council in 1953 (following a smaller, preliminary meeting in 1951), *psycholinguistics* as a term and an endeavor was rediscovered, and an ambitious agenda for cooperative research was published (Osgood & Sebeok, 1954). The specific linguistic and psychological theories that the participants brought to the table look dated now, as do many of the information theory concepts they embraced. By contrast, the research goals and methods were carried forward with surprisingly little change into the next era of psycholinguistic research (section 2.2). For example, the goal of establishing the *psychological reality* of linguistic constructs, such as the phoneme, was to be pursued by such methods as the analysis of speech errors. A more ambitious idea was to "suppose that speech sounds can be regarded as occupying positions in a multidimensional space" (p. 78) and to discover that space by applying an advanced psychological scaling technique, *factor analysis*, to subjects' judgments of the similarity between particular sounds. One participant adapted this method to semantics, uncovering three dimensions of connotative meaning (evaluation, potency, and activity), and obtained cross-cultural evidence of their universality (Osgood et al., 1957). The title of this book, *The Measurement of Meaning*, says it all: for latter-day Hullian psychologists, it was OK to study meaning if you could come up with a good way to measure it.

The idea of a space of speech sounds was linked (in ways too complex to pursue here) to Roman Jakobson's analysis of the world's phonemes in terms of a small set of binary *distinctive features* (e.g., *voiced* vs *unvoiced*, *nasal* vs *oral*, *front* vs *back*). A variety of the distinctive feature approach called *componential analysis* also found application within anthropological linguistics. For example, seminar participant Floyd G. Lounsbury (1956) analyzed kinship terms using semantic features like *female/male kinsman*, *female/male ego*, *generation 1/generation 2*, as well as some less obvious ones.

No one suspected it at the time, but the rug was about to pulled out from under these structural linguists and their dance partners from psychology; the youngest and most nimble would learn to dance to a new linguistic tune, and the others would end up on the sidelines. The change began with a post-Bloomfieldian who was unusually adept at discovering discovery procedures – Zellig Harris at the University of Pennsylvania. He and other linguists had achieved considerable success with phonology and morphology, but in the early 1950s syntactic analysis was still underdeveloped. In the quest to make syntax tractable, Harris had the idea of normalizing complex sentences by using *transformations* to relate them to simpler *kernel sentences*; further analysis could then focus on the kernels. For example, the passive sentence "Titchener was defeated by the behaviorists" and the cleft sentence "It was the behaviorists who defeated Titchener" can both be transformationally related to the same kernel sentence: "The behaviorists defeated Titchener."

The notion of a transformation blossomed into a revolution in the hands of Harris's student Noam Chomsky. After receiving his M.A. in 1951, Chomsky moved to Harvard as a Junior Fellow of the Society of Fellows from 1951 to 1955. During this period he wrote a large, difficult work entitled *The Logical Structure of Linguistic Theory*, which remained unpublished until 1975. He submitted part of it for his Ph.D. from the University of Pennsylvania in 1955 and joined the faculty at MIT (home of the aforementioned Building 20). His first book, *Syntactic Structures* (1957), made Chomsky's ideas more accessible. Taking a combative stance, Chomsky won only a few converts from the established generation of post-Bloomfieldian linguists. But, more important, he and his collaborator Morris Halle attracted the best of the new generation of graduate students to MIT, and it was these students who were hired as linguistics departments were begun or expanded in the 1960s.

One of Chomsky's key departures from Bloomfieldian linguistics was that he construed a grammar as a generative system – a set of rules that would generate all and only members of the infinite set of grammatically well-formed sentences of a language. Chomsky took up the question of what sort of computational system (automaton) was needed to realize a generative grammar for natural language (cf. Harris's search for a computational system that could discover a grammar from a corpus of sentences). Chomsky argued that two sorts of systems then being considered were inadequate. One was a finite state automaton (Markov process), which consists of a finite number of states and probabilistic transitions between states. As applied to generating sentences, in an initial state there would be a choice of words with which to begin the sentence (e.g., *Noam* or *Linguists*); selecting one of them determines the next state, where again there is a choice of words (e.g., the choices might include *proposes, disposes,* and *sleeps* following *Noam* versus *propound, disagree,* and *sleep* following *Linguists*). An important limitation is that such a device has no memory of the path by which it reached its current state; it has access only to the next set of choices (the possible transitions from that state to the next). Chomsky argued persuasively that no finite state automaton could be adequate to the task, because English (as an example of a natural language) is not a finite state language. The complete argument cannot be easily summarized (it includes the assumption that natural languages have an infinite number of sentences, at which some readers balked), but the flavor is conveyed by considering the difficulty of designing a finite state automaton to generate sentences with embedded clauses ("The woman who was talking to the astronauts is the President") while maintaining the correct dependencies between nonadjacent words (e.g., subject–verb agreement between *woman* and *is*).

Next, Chomsky argued for the inadequacy of phrase structure grammars, which generate phrase structure trees by successively applying rewrite rules such as $S \rightarrow NP\ VP$ and $NP \rightarrow Adj\ N$. (These rules state that a sentence can be composed of a noun phrase followed by a verb phrase; in turn, the noun phrase can be composed of an adjective followed by a noun. This STRUCTURAL ANALYSIS is often displayed in an inverted tree structure diagram, with S at the top.) He pointed out that such grammars would have to be unnecessarily complex due to their inability to take advantage of such regularities as (a) the underlying similarity between a kernel sentence and sentences related to it transformationally, and (b) the dissimilarity between different readings of certain ambiguous sentences, such as "Flying planes can be dangerous."

Accordingly, Chomsky advocated transformational grammars, in which rewrite rules generate an underlying tree structure, and transformational rules then apply to obtain a derived phrase structure. In our example, an active sentence and a corresponding passive sentence would have the same initial phrase structure (called *deep structure*) but different derived structures (called *surface structure*); the passive sentence would have the passive transformation included as part of its derivation. Chomsky worked out a fragment of a transformational grammar of English in *Syntactic Structures* and added revisions and elaborations in what came to be known as his Standard Theory in *Aspects of the Theory of Syntax* (1965). Over a long career, he has repeatedly revisited the question of what kind of grammar best captures generalizations about language (see Article 15, LINGUISTIC THEORY).

In many respects, Chomsky's transformational grammar was a natural extension of the structuralist program of Bloomfield and Harris. Like the structuralists, he focused on the formal structures of a language. His claim of the autonomy of syntax – the claim that we can model syntactic knowledge independently of concerns for meaning (semantics) or the pragmatics of communication – comported well with structuralist principles. But Chomsky incorporated his grammar within an ambitious theoretical framework that repudiated the behaviorism of the structuralists by emphasizing the creative open-endedness of syntax and especially by moving to a Cartesian mentalism in the 1960s (see section 2.2).

1.5 Inside the delivery room: the events of 1956

So far we have emphasized the intellectual developments which provided the foundation for cognitive science. We have also seen that figures in the history of cognitive science often did not work alone. One of the important forums in which they communicated were conferences.

Within established fields of inquiry, such as cognitive science has now become, the major conferences are highly formalized. They occur on a regular basis, and although each year a different program committee places its own stamp on the gathering by emphasizing some topics and de-emphasizing others, the programs usually adhere to an established script. At the outset of a new field of inquiry, however, the meetings have a much different character. Investigators are putting forward programs, not incremental advances. The discussion from the audience focuses on the wisdom of the new program and is not consumed with challenges directed at details. Informal conversations in the hallways or over drinks are even livelier than in more settled times. Furthermore, a new scientific inquiry does not draw upon an already established lineage of investigators and may even rely on the ability of investigators from different disciplines to communicate across differences in terminology, techniques, and criteria for success. Sometimes the salience of what investigators in other disciplines are doing is so tangible that it is possible, at least momentarily, to glimpse the value of an integrated pursuit and transcend the differences.

We have noted a number of conferences already in which such crossing of disciplinary boundaries occurred, but a meeting at MIT on September 10–12, 1956, seems to have taken this a step further, so much so that George Miller fixes on the second day of the conference, September 11, as the birthdate of cognitive science ("the day that cognitive science burst from the womb of cybernetics and became a recognizable, interdisciplinary adventure in its own right" (Miller, 1979, p. 4)). Miller reports going away from the 1956 Symposium on Information Theory "with a strong conviction, more intuitive than rational, that human experimental psychology, theoretical linguistics, and the computer simulation of cognitive processes were all pieces from a larger whole, and that the future would see a progressive elaboration and coordination of their shared concerns" (p. 9).

The papers in this symposium brought together some of the major contributions discussed earlier. The first day was devoted to coding theory and included papers by Shannon among others. A symposium on automata started the second morning, and the first paper, "The logic theory machine: a complex information processing system" by Newell and Simon presented their Logic Theorist's proof of theorem 2.01 of Whitehead and Russell's *Principia Mathematica*. The following paper, by Rochester, Holland, Haibt, and Duda, presented their computer implementation of Hebb's neurophysiological theory of cell assemblies. The next symposium, on information sources, included a paper by the young Chomsky entitled "Three models of language," in which he presented his arguments for transformational grammar. The third symposium of the day, on information users, included Miller's paper on the magic number 7.

One reason for focusing on this conference as representing the birth of cognitive science is that it did more than bring together some of the most important accomplishments of cognitive science's gestation period. It also made clear that the basic theme linking the various disciplines of the mind was a common conception of the mind as engaged in processing information. Shannon had formulated the idea of information as a measure of what could be transmitted over a channel like a telephone wire, but the different fields began to focus more on how information could be represented and operated upon. From linguistics, Chomsky presented the idea that the representation of linguistic knowledge involves rewrite rules and transformational rules. From psychology, Miller's discussion of the magic number 7 introduced the notion of memory as a limited-capacity information storage system that forces us to use hierarchical encodings of information. Rochester and colleagues' implementation of Hebb's model of cell assemblies suggested how information processing might be accomplished in the brain as well as in computers. Thus, the various speakers found that they shared the conception of the mind as an information processing system and saw that different disciplines could each contribute to understanding it.

The year 1956 stands out as seminal beyond these important papers – we have encountered it a number of times. New twists were added to some traditional fields, including Stevens's magnitude estimation technique in psychophysics. Other fields became more explicitly cognitive: 1956 was the year that Bruner, Goodnow, and Austin published *A Study of Thinking* and Festinger published his cognitive dissonance theory. The Dartmouth Conference got artificial intelligence off to a start. Anthropologists Goodenough and Lounsbury each published a componential analysis of kinship terms. Ulric Neisser received his Ph.D. from Harvard, and within a few years so did many others of the new generation of researchers who would bring cognitive science to initial maturation. We now turn to that part of the story.

2 Maturation, 1960–1985

A birth represents a transition between a time of gestation and a period of maturation. If we assume that cognitive science was born in the late 1950s (Miller's Tale of 1956 plus or minus), the next 25 years, roughly from 1960 to 1985, represents its initial maturation. During this period cognitive science began to be refined and to bear fruit, both intellectually and institutionally.

In this section we will examine a number of these intellectual accomplishments, including the building of new institutional frameworks in cognitive science. One consequence for refinement, largely unintended, was that some elements of the original broad perspective of cognitive science were initially left out. These included, most notably, neuroscience and artificial neural networks. Kintsch, Miller, and Polson, in their 1984 collection *Method and Tactics in Cognitive Science*, for example, included only AI, linguistics, and psychology as cognitive science disciplines. The reasons for this shrinking in perspective will be discussed both in this section and in section 3 (where we consider how fields like neuroscience were reincorporated into cognitive science after 1985). But we shall also see that another discipline, philosophy, started to play a more important role during the phase of initial maturation.

2.1 *Early development: a distinctively cognitive model of mind*

During the gestation period the different disciplines of cognitive science each developed proposals for thinking about cognitive processes without managing to produce a cohesive model of the mechanisms that might realize cognition. But in 1960 George Miller, together with Eugene Galanter and Karl Pribram, developed just such a proposal for a basic cognitive mechanism, which they called a *TOTE* unit. They described these units and argued that they could provide the basis for intelligent behavior in *Plans and the Structure of Behavior*, perhaps the most influential of the early books in cognitive psychology. *Plans* was composed during the 1959–60 academic year, which Miller, Galanter, and Pribram spent at the Center for Advanced Studies in the Behavioral Sciences at Stanford. Like Miller, Galanter had been trained in psychophysics by Stevens; however, he also had some background in mathematical psychology, engineering, and philosophy (especially logic) and so was attracted to the idea that algorithmic procedures might provide an account of mental activities. Pribram, a student of Karl Lashley, was then a young neuroscientist.

During the year at the Center, Miller organized a conference at which there was a lively debate between experimentalists and mathematical modelers. According to Galanter, the experimentalists criticized the mathematical modelers for failing to provide an explanatory mechanism for the functional relationships between variables and behavior which their equations described. For example, in a simple linear learning model, the probability of a correct response on the next trial equals the probability of a correct response on the current trial plus an increment that is proportional to the amount still to be learned. How high a proportion may depend upon such factors as the size of the reward. But what internal mechanism would work in such a way as to produce a proportional increment? And beyond that, what mechanisms might be in place to produce more complex and interesting behaviors? *Plans* was envisaged as a speculative book that would help chart a new course: "The aim, we all agreed, was to replace behaviorism (really our name for associationism) and trivial experimentation (the T-maze and the Skinner box) with the new insights and maybe even the new technology from computational logic, linguistics, and neurophysiology" (Galanter, 1988, p. 38).

Miller, Galanter, and Pribram took as their focus purposive action (executed *Plans*) and proposed that Plans took the form of TOTE units: Test–Operate–Test–Exit. Borrowed from the feedback loops of cybernetics, TOTE was advanced as an alternative to the classical reflex arc as a model for the basic unit of mental activity. Galanter relates the discussion which gave rise to the TOTE unit:

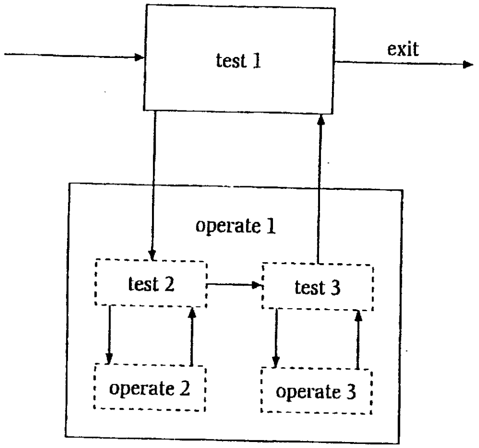

Figure I.6 A recursively embedded set of TOTE units as proposed by Miller, Galanter, and Pribram (1960). If test 1 reveals that its goal was not satisfied, operation 1 is invoked, which itself imposes tests and operations. The operations specified by any test would be repeated until the test showed that the goal was satisfied, at which point that TOTE unit would be exited.

> At one point, George proposed that we examine some intentional human act.
> "Flying a plane," I suggested.
> "No – too much. How about crossing a street. An equally dangerous act in the Bay area," Karl responded. I went to the blackboard and started a flow chart. The boxes, lines, and arrows snaked around the board as step after step was drawn.
> "No," George said, "all that stuff on the board is only a string of reentrant reflexes. Let a whole piece of the action be repeated until it's finished."
> "How will it know?" from Karl.
> "With a cybernetic test," replied George.
> "But how do I draw it?" I asked.
> "Like this," said George, and the TOTE replacement for the reflex was designed.
> (p. 40)

The key idea was that within an agent an initial test would evaluate a current situation against a goal; if they did not match, an operation was carried out to reduce the difference. The cycle of test and operate would be repeated until the goal was matched, at which point the TOTE unit would be exited. Miller, Galanter, and Pribram also proposed that TOTE units could be recursively organized by inserting TOTEs within the operation cycle of another TOTE, as shown in figure I.6. One of the applications they envisioned (but did not work out in detail) was a realization of Chomsky's phrase structure trees and transformations.

Historically, the TOTE unit can be viewed as realizing Lashley's (1951) call for some sort of central, hierarchically structured control over sequential behavior (he had argued persuasively that fluent sequences such as those in typing and speaking occurred too rapidly to be produced by simple associative chains from one element of the sequence to the next). The realization was only schematic, but traces of some of the ideas (considerably modified) can be seen in Newell and Simon's production systems architecture (see Article 42, PRODUCTION SYSTEMS).

2.2 *Learning to talk: Chomsky's impact reaches psycholinguistics*

One of the first sustained collaborations in cognitive science was that between psychology and linguistics. As we saw in section 1.4, the collaboration began in the early 1950s, prior to the cognitive turn in both disciplines. In 1957 Chomsky proposed his transformational grammar in *Syntactic Structures*. The impact began to be absorbed within linguistics, then psycholinguistics, and then helped to shape the new cognitive psychology as well.

In the meantime, Chomsky himself was elaborating his initial proposals. He made a number of changes to the grammar itself, but perhaps more important was his movement towards a Cartesian rationalism and bold claims about *Language and Mind* (Chomsky, 1968). Adopting a strongly mentalistic perspective, Chomsky proposed that (a) generative grammar was a formalized expression of people's *tacit knowledge* of their language, and (b) the primary data for linguistic analysis should be speakers' judgments as to which sentences are grammatical. Chomsky was well aware that people sometimes produce ungrammatical sentences. However, he introduced a distinction between *competence* and *performance*, which enabled him to attribute to people correct tacit knowledge of their language in the face of flawed performance (e.g., a speaker might forget that the subject of a sentence in progress was singular, and hence fail to inflect the verb with the agreement affix, -*s*).

Early in the development of these ideas, Chomsky (1959) repudiated behaviorism in a review in *Language* of B. F. Skinner's 1957 behaviorist account of language, *Verbal Behavior*. In part of the review Chomsky focused on the vacuity of Skinner's major explanatory concepts – stimulus, response, and reinforcement – in explaining language use. For example, the stimulus control for a given behavior might be defined well enough in animal conditioning contexts, but not in the context of an adult human using someone's name (we do not use a name only in the presence of the person whose name it is). Chomsky also emphasized that language use is a creative activity, in the sense that there is no bound to the novel but grammatically well-formed sentences one might produce or hear. Finally, Chomsky introduced a nascent form of what later came to be known as his *poverty of the stimulus* argument for the innateness of basic linguistic competence: "The fact that all normal children acquire essentially comparable grammars of great complexity with remarkable rapidity suggests that human beings are somehow specially designed to do this" (p. 58). As he later developed the idea (e.g., Chomsky, 1965, 1968), the sentences in a child's environment provide too impoverished a database to make it credible that ordinary learning processes can account for the child's competence; an innate *Universal Grammar (UG)* must guide the child's inductions from input (see Article 45, INNATE KNOWLEDGE).

Chomsky's incursions into psychology generated considerable controversy, even to this day. But his formalizations of generative grammar have probably had a broader practical impact in psychology, because they suggested how mental representations might look at a time when psychologists were ready to ask this question. George Miller was ready earlier than most. His interest in information theory had led him to explore the potential of statistical approaches for understanding human information processing.

Although his goal had not been to provide a grammar for English, Miller quickly appreciated the importance of Chomsky's contention about the kind of automaton needed to generate natural human languages and saw its relevance to his own project: if finite state automata were inadequate, Miller's own statistical approach was doomed. He employed Chomsky as his assistant for a summer seminar on mathematical psychology in 1957, and together they wrote a 1958 paper on finite state languages.

In a more practical vein, this led to a research program for Miller in which cognitive activity was construed in terms of operations on symbolic representations. Chomsky's 1957 grammar had transformations derive one tree structure from another. While it is certainly not necessary to use a linguistic grammar as a direct model for positing a series of mental operations performed by a speaker or hearer of a language (it could be viewed as simply a compact representation of the grammatically well-structured sentences of a language), the idea was attractive enough to launch an exciting era of psycholinguistic research. Using Chomsky's transformational grammar to suggest experiments (rather than speculations about the uses of Plans as in the 1960 book), Miller gathered collaborators and presented the preliminary results of several studies in the *American Psychologist* (Miller, 1962). Both memory for sentences (Mehler) and response times (Miller, McKean and Slobin) revealed that the more transformations in a sentence's derivation, the more difficult it was to process. This was taken as evidence for transformational grammar's *psychological reality*. (The phrase was borrowed from Edward Sapir, but one goal was to supersede the psychological reality claims of the psycholinguistics of the 1950s.)

The specific form of this research program in psycholinguistics was short-lived, as it soon became apparent that not all transformations added processing time. The new generation of psycholinguistic researchers (particularly Jerry Fodor, Merrill Garrett, and Thomas Bever at MIT) concluded that the relation between competence and performance must be more abstract than originally thought. Also favoring this more nuanced approach was the fact that Chomsky changed his theory over the years so as to de-emphasize transformations, removing the explanation for the data that Miller and others had gathered. The focus of psycholinguistic research shifted from transformation counting to explorations of the psychological reality of deep structure and of the processes involved in parsing surface structures, and became broader still as the field matured (see Article 14, LANGUAGE PROCESSING).

The heady psycholinguistics of the 1960s eventually settled into a more sedate *normal science* mode, but cognitive psychology and cognitive science more generally had been permanently transformed by Chomsky's ideas about how to describe and explain the linguistic competence of speakers. To get some feel for the extent of the change, consider the titles of articles published in the *Journal of Verbal Learning and Verbal Behavior*. The first volume appeared in the same year as Miller's paper on transformations (1962) and did not yet reflect Chomsky's influence; among the titles were "Verbal mediation in paired-associate and serial learning," "Aural paired-associate learning: pronunciability and the interval between stimulus and response," and "Associative indices as measures of word relatedness: a summary and comparison of ten methods." A few papers reflecting Chomsky's influence appeared in 1963, and by 1968 there were a number of papers with such titles as "The role of syntactic structure in the recall of English nominalizations," "The perception of grammatical relations in sentences: a methodological exploration," and "Semantic distinctions and memory for complex sentences." Finally the title of the journal caught up with the content in 1985, when it was renamed *Journal of Memory and Language*.

Chomsky's impact was not limited to psycholinguistic studies of adult subjects in laboratories. People investigating children's language began calling their field *developmental psycholinguistics* and using Chomskian grammars as a framework. Eric Lenneberg published his landmark *Biological Foundations of Language* in 1967, helping to make *neurolinguistics* an exception to the general neglect of neuroscience in cognitive psychology during the 1960–85 period. *Computational linguistics* became a distinct research area, in which the design of parsing programs was one focus. Even *sociolinguistics* and *anthropological linguistics* made more room than usual for formal approaches to language. These fields are discussed in several of the chapters that follow; here we must return to developments in psychology.

2.3 *A first home: the Center for Cognitive Studies at Harvard*

As noted earlier, two of the founders of today's cognitive psychology, Bruner and Miller, were housed in separate departments at Harvard. The difference in orientation between these two departments was considerable. Yet, Bruner and Miller clearly recognized their intellectual affinities and had been teaching a course called "Cognitive Processes" together for several years. In 1960, after Miller's return from the Center for Advanced Studies in the Behavioral Sciences, they put together a proposal for a Center for Cognitive Studies. It was funded for ten years by the Carnegie Corporation. Although Bruner and Miller were both psychologists, the Center was designed from the outset to have an interdisciplinary character.

One of the things the Center did was to provide support for long-term research fellows, many of whom were scholars at the beginning of their careers who would go on to make their own mark in such cognitive science fields as COGNITIVE AND LINGUISTIC DEVELOPMENT, CONCEPTUAL ORGANIZATION, IMAGERY AND SPATIAL REPRESENTATION, LANGUAGE PROCESSING, MEMORY, and ATTENTION. Among them were Arthur Blumenthal, Janellen Huttenlocher, Paul Kolers, David McNeill, Donald Norman, and Nancy Waugh. These fellows developed a large number of research projects, many of them in collaboration with one another, with Miller or Bruner, or with one of the annual visitors. There were also a number of graduate students in the Harvard departments who worked with the research fellows and visitors and later had influential careers in cognitive science, such as Ursula Bellugi, Susan Carey, Patricia Greenfield, Jacques Mehler, and Dan Slobin.

One of the major justifications for the Center was to bring in outside visitors to spend a year doing research and contributing to the community there. As Miller relates, "The way this worked was that he [Bruner] and I got together at least once each year over a bottle of madeira (or was it a bottle of port?) and discussed people whose ideas we found exciting. Anyone whose ideas appealed to both of us was invited to join us for a year" (Miller, 1979, p. 11). A number of scientists, both from the Cambridge area and the rest of the world, spent a year at the Center. Among these were researchers focusing on language, such as Chomsky, Roger Brown, Roman Jakobson, Jerrold Katz, and Willem J. M. Levelt. Roger Brown's longitudinal study of the early language of three children provided the basis for much subsequent work in developmental psycholinguistics (see Article 6, COGNITIVE AND LINGUISTIC DEVELOPMENT). Levelt went on to establish the Max-Planck-Institut für Psycholinguistik at Nijmegen, one of the premier centers in the world for psycholinguistic research (see Articles 14, 24, and 25: LANGUAGE PROCESSING, UNDERSTANDING TEXTS, and WORD MEANING). Several other visitors, such as Daniel Kahneman, Amos Tversky, and Peter Wason, made REASONING and PROBLEM SOLVING a focus of their research. Each of them produced evidence that human reasoning does not conform to the norms of proper reasoning advanced in logic (see Article 44, HEURISTICS AND SATISFICING).

A major activity of the Center was its Thursday afternoon colloquia, a series which was announced to the broader public and drew audiences from a number of other institutions. This series continued a previous series that Bruner had established in 1952 as part of his Cognition Project (the Center's predecessor). To give the flavor of the breadth of topics covered in the series, the schedule for the first year is reproduced in box I.1.

While formal activities such as colloquia are perhaps the easiest to document, participants at the Center tended to focus on the informal activities. Thus Norman and Levelt relate:

> Bruner states that "the intellectual life of the Center revolved around the seminars, the Thursday lunches, and the weekly colloquia." Perhaps. But that is not our memory. For us, the intellectual life was in the routine daily activities, in the offices and halls, in the labs late at night, and in the social interactions. The excitement was in the personal interaction and the private discussions and arguments. The formal seminars and lunches and colloquia were, well, formalities: the public display of the refinements. . . .
>
> The prototypical lunchtime seminar – or at least, prototypical in our memory – is of everyone assembled around the large wooden seminar table with an active, young cast of protagonists (perhaps Mehler, Bever, Fodor, and Katz), each paraphrasing and explaining to the lunchtime audience what the one had tried to explain to us what another had just said what yet another had just previously said that Noam would have said in retort to whatever the issue was at the time. All the time, Chomsky sitting and listening to the others explaining his mind. (Norman and Levelt, 1988, pp. 100–1).

The Harvard Center provided a prototype for many cognitive science centers developed later and left a lasting imprint, but after a decade it was dissolved. Its core consisted of two individuals, but in 1967 Miller left for Rockefeller University. Bruner directed the Center by himself for a couple of years, but it was closed in 1970, and Bruner moved to Oxford University in 1972.

2.4 *Cognitive psychology learns to walk and travels to other institutions*

Although Bruner, Miller, and those connected with the Center at Harvard had played an important role in giving a cognitive focus to psychology, the endeavor quickly became mobile, and cognitive psychology became part of the program in experimental psychology at numerous institutions.

The new endeavors became known as *cognitive psychology*, largely due to Ulric Neisser's influential book in 1967 by that name. In addition to a name, this book gave cognitive psychology a broad vision. Neisser, then a very junior professor at Brandeis and frequently associated with the Harvard Center, had entered psychology just as the crucial ideas of information theory and cybernetics were being adopted by Miller and his associates. He describes taking Miller's course, "The Psychology of Speech and Communication," in 1949: "George taught us about decibels and filters, phonemes and morphemes, Shannon's theorem and Zipf's law, coding principles and the redundancy of English. Naturally there was no standard textbook for such a course, so he was writing one himself. In 1949, we worked from mimeographed copies of the first draft. We didn't know then that it was a first draft of the future" (Neisser, 1988, p. 82).

Box I.1 Thursday afternoon colloquia, Harvard Center for Cognitive Studies, 1960–1

November 10: Jerome S. Bruner, Harvard University, "Similarity and Difference: Two Approaches to Knowing"

November 17: Ernst Mayr, Harvard University, "A Discussion of the Phylogenetics of Information Processing"

December 1: Daniel Berlyne, Boston University, "Epistemic Curiosity"

December 8: Roger Brown, Massachusetts Institute of Technology, "The Acquisition of Grammar"

December 13: Roman Jakobson, Harvard University, "Infants' and Aphasics' Testimonies on Language and Thought"

January 12: Walter A. Rosenblith, Massachusetts Institute of Technology, "Computer-aided Electrophysiological Studies in Sensory Communication"

January 19: Richard Held, Brandeis University, "Exposure to Strange Environments and an Ordering Principle in Perception"

January 26: Raymond A. Bauer, Harvard University, "A New Look at an Old Problem: Rational Versus Emotional Appeals"

February 2: Eric H. Lenneberg, Harvard Medical School, "The Biological Matrix of Speech"

February 9: Zoltan Dienes, University of Adelaide, "A Theory of Learning Mathematics"

February 16: Frederick C. Frick, Massachusetts Institute of Technology, "Pattern Recognition"

February 23: Ulric Neisser, Brandeis University, "A Theory of Intuitive Thinking"

March 2: Gordon Allport, Harvard University, "Intuition Revisited"

March 9: David Page, University of Illinois, "On Teaching Mathematics"

March 16: Walter Mischel, Harvard University, "Cognitive Activity and Delay of Gratification"

March 23: Ivor A. Richards, Harvard University, "Film Sequences in the Investigation of Learning"

March 30 and April 13: Noam Chomsky, Massachusetts Institute of Technology, "Grammatical Factors in the Perception of Sentences"

April 20: Harold Conklin, Columbia University, "The Cultural Relevance of Cognitive Contrast"

April 27: Jules Henry, Washington University, "Cultural Factors in Elementary School Readers"

May 4: Gerard Salton, Harvard University, "Some Problems in Automatic Information Retrieval"

May 11: David French, Reed College, "Anthropology, Ethnoscience, and Cognition"

Even then, however, Neisser's interests were broader than those Miller was pursuing; he had sympathies with the Gestalt psychologists, many of whom had just emigrated to the USA, and he accordingly began graduate school at Swarthmore with Köhler. But after a couple of years Neisser returned to work with Miller and then, like several other pioneers, completed his Ph.D. in psychophysics under Stevens. While

Neisser found aspects of information theory attractive, he was restless for experiments that could genuinely reveal the nature of cognition. At an MIT symposium on information theory he met Oliver Selfridge, then working on his Pandemonium model of pattern recognition (see section 1.1.2). Neisser and Selfridge collaborated on visual scanning experiments and on further development of the Pandemonium model. From these and other experiences, Neisser developed a vision of what cognition might be like:

> By 1964, it had come together in my head. In principle, I thought, one could follow the information inward from its first encounter with the sense organ all the way to its storage and eventual reconstruction in memory. The early stages of processing were necessarily wholistic (an idea I borrowed from Gestalt psychology) and the later ones were based on repeated recoding (an idea borrowed, even more obviously, from George Miller). But the processing sequence was by no means fixed; at every point there was room for choice, strategy, executive routines; individual constructive activity. Noam Chomsky's linguistic arguments had shown that an activity could be rule governed and yet indefinitely free and creative. People were not much like computers (I had already sketched out some of the differences in a 1963 *Science* paper), but nevertheless the computer had made a crucial contribution to psychology: It had given us a new definition of our subject matter, a new set of metaphors, and a new assurance. (Neisser, 1988, p. 86)

With this vision in mind, Neisser took a leave from Brandeis and wrote *Cognitive Psychology*. Published in 1967, it served to both introduce and synthesize the work on information processing that was beginning to burgeon, particularly emphasizing attention and pattern recognition, and it quickly became the bible for a new generation of students.

One set of topics was rooted in Broadbent's (1958) argument that there are three types of MEMORY stores. Oversimplifying and using a later set of names for the stores, here are some highlights. (1) Sperling (1960) found evidence for the first type, a *sensory register* in which a visual or auditory stimulus briefly persists after it has been turned off. Sperling devised a way to actually measure the rapid decay of information from this register. Presented with 3 by 3 array of letters for 50 msec, subjects can generally report four or five letters (i.e., one or two per row). But if one row is cued by a tone exactly when the array is turned off, subjects can still "see" it briefly and, by concentrating just on their image of that row, report all three letters. If the cue is delayed for a split second, the image in the sensory store has partly decayed, and the advantage of cueing is less. At a delay of one second, decay is complete. (2) The second type of memory discussed by Neisser is the *short-term store* that was the focus of Miller's seven-plus-or-minus-two paper. Peterson and Peterson (1959) traced decay from this store by presenting three consonants to subjects (well within the memory span) and then requiring them to count backwards by threes for periods varying from 3 to 18 seconds before recall. Ability to recall the consonants dropped sharply between 3 and 9 seconds' delay. In an alternative procedure, Waugh and Norman (1965) used a sequential probe technique to prevent rehearsal. Subjects heard 15 digits and then, probed with one of the digits, tried to report the digit that had followed it. The further back the probe had been in the sequence, the poorer the performance. Waugh and Norman attributed this to later digits replacing earlier digits in the limited number of slots in the short-term store. (3) The third type of memory discussed by Neisser is a *long-term store*. Waugh and Norman argued that when subjects have time to rehearse an item, it thereby retains its place in the short-term store and also gains a chance of being added to the long-term store. In general, the short-term store has an auditory character (the inner voice), whereas the long-term store is meaning-based. (Later theories, such as that of Alan D. Baddeley (1976), include a visual short-term store as well.)

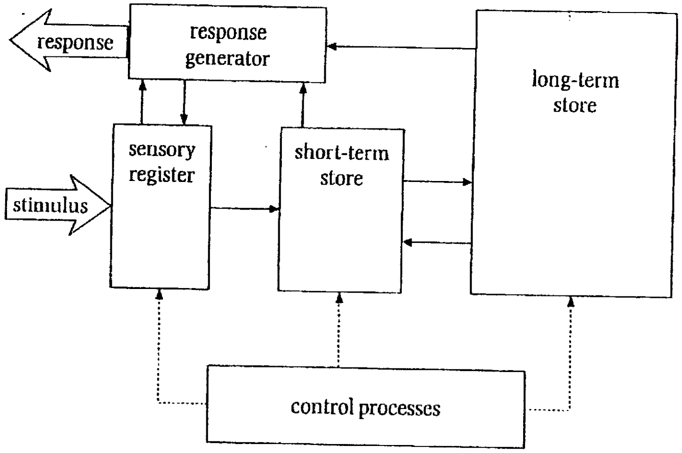

Figure I.7 Atkinson and Shiffrin's (1968) memory model.

This idea of a short-term versus a long-term store was not new. Waugh and Norman cited William James on the two memory systems and used his terms of *primary* and *secondary memory* to refer to them. What they added were an experimental technique, a theory about rehearsal, and a mathematical model in which the probabilities of retrieval from the two stores were statistically independent. These contributions played an important role in the many information processing models of memory that emerged around the same time that Neisser's book was published. The most influential integrative model (shown in figure I.7) was developed at Stanford University by Richard Atkinson and Richard Shiffrin (1968). Key to their model was a distinction between fixed and flexible structures. The sensory register and short-term and long-term stores are fixed. The flexible structures are control processes which operate on the fixed structures. For example, ATTENTION is a control process that determines which information gets transferred from the sensory register to the short-term store (in a new format), and rehearsal is a control process that determines which information is retained there long enough to be copied into the long-term store. Items could be lost from the short-term store (bumped by new items or simply through decay), but Atkinson and Shiffrin claimed unlimited capacity for the long-term store. Forgetting would be due to retrieval failure, not loss. An important part of their work was to develop mathematical models of the workings of this system that would specify, for example, the probability that an item would be retrieved from at least one store during a memory test.

These are just a few pieces of work from a very active period. One way to view the spread of the cognitive approach is to focus on how it developed at different universities. We have picked three institutions to get a representative range of developments – each gained strength in cognitive psychology by a different route. A portrait of each of these exemplars is suggestive of the important developments occurring at a number of other institutions as well.

2.4.1 *Stanford University*

Stanford had a well-established and well-regarded psychology department that was nimble enough to quickly gain strength in cognitive psychology in the 1960s. It attracted a stellar faculty, including Atkinson, Bower, and Shepard, and has probably been the premier university in producing graduate students who went on to shape the field. Two younger faculty members whose influential work helped to shape the department in the 1970s were Herbert Clark and Edward Smith. Clark's work on LANGUAGE PROCESSING was strongly interdisciplinary; he often collaborated with Eve Clark in the linguistics department. The Clarks were proponents of semantic feature approaches (descended from Jakobson's phonemic features) and produced models of sentence processing which specified, for example, the extra time taken to process negative sentences or to deal with mismatches between word order and referent order. Smith's research was on CONCEPTUAL ORGANIZATION. With graduate students Lance Rips and Edward Shoben, he predicted reaction times from a model of semantic memory that used multidimensional scaling to represent basic concepts like *robin* and superordinates like *bird*. Multidimensional scaling, developed more recently than the factor analysis used by Osgood, is a family of methods for using data on pairwise item comparisons to uncover a particular representation of those items – a multidimensional psychological space. For example, Rips, Shoben and Smith (1973) found birds to be organized in a space with dimensions they labeled as size and predacity. In general, greater distances between the point representing a prototypical bird and the point representing a particular bird were associated with longer reaction times: it took longer to verify "A duck is a bird" than "A robin is a bird."

Shepard's forte has been the use of mathematical models to understand human thought. His core concern has been to show that psychological laws can be obtained by formulating them in relation to a psychological space. His formal work in this vein (e.g., on nonparametric multidimensional scaling in the 1960s and on generalization later) is highly respected and influential, but the work that has received the broadest attention is on mental imagery and mental rotation. In a study with his colleague Jacqueline Metzler, he presented subjects with pairs of geometrical forms. Subjects were asked whether they were images of the same object (differing only in rotation, versus mirror images). Reaction times increased linearly with the degree of rotation, which Shepard and Metzler (1971) interpreted as evidence that subjects were mentally rotating one of the figures at a constant rate. These initial results inspired further studies by Shepard and his students, especially Lynn Cooper (1975). After so many years of anti-mentalism, and the recent emergence of a proposition-based mentalism, these results persuaded many that analog images and processes not only existed but could be brought under experimental control and measurement (see Article 12, IMAGERY AND SPATIAL REPRESENTATION).

Gordon Bower's early work was on mathematical models of memory; like Shepard, he moved towards more cognitively oriented research on the nature of mental representation. En route he did work in the 1960s on the role of organization, mental imagery, and mnemonics in memory. A researcher of unusual breadth and legendary as a mentor, he directed the dissertations of numerous students in the 1960s through 1980s who would become major contributors to almost every major area of cognitive science, including investigations of both analog and propositional mental representations. To provide just a few examples from the early 1970s: on the analog side, Stephen Kosslyn designed ingenious experiments to support his claim that the mind employs image-like REPRESENTATIONS; he is now a major figure in cognitive neuroscience (see section 3.2). Preferring a propositional approach, John Anderson constructed a semantic network-based model of associative memory (HAM) and performed numerous experiments to test its predictions. The earliest work is described in Anderson and Bower's 1973 book, *Human Associative Memory*. Much of Anderson's subsequent career, based at Carnegie–Mellon University after a few years at Yale, has been devoted to developing a computational framework, ACT* (Adaptive Control of Thought), which incorporates both PRODUCTION SYSTEMS and semantic networks that allow spreading activation between nodes in memory (Anderson, 1983). Another influential Bower student is Robert Sternberg, whose studies of mental representation and processing as a graduate student launched him into a career focused on new approaches to intelligence. Sternberg's concerns have included decomposing intelligence test tasks into their cognitive components, broadening the conception of intelligence beyond academic intelligence, and understanding individual differences.

Students and faculty in psychology at Stanford had ample opportunity to benefit from a talented faculty in areas of psychology besides cognition and also from groundbreaking work outside their own department. Particularly noteworthy were the AI researchers in the computer science department (John McCarthy, Edward Feigenbaum, and Roger Schank) and nearby research centers (e.g., Daniel Bobrow at Xerox Palo Alto Research Center and Bertram Raphael at Stanford Research Institute). Also, the Center for Advanced Study in the Behavioral Sciences brought many of the world's leading cognitive psychologists to Stanford as one-year visitors.

2.4.2 *University of California, San Diego (UCSD)*

In contrast to Stanford, UCSD was a new university that admitted its first undergraduates in 1964 and initially emphasized the sciences. It attracted a vigorous faculty that welcomed the opportunity to design programs from the ground up, and hence provided a relatively open niche to be colonized by cognitivists. The first three faculty arrived in 1965, including George Mandler as chair. Like Bower, he is renowned as a mentor and contributed some of the early studies that persuaded psychologists that memory involves active organizational processes. He also encouraged the adaptation of Signal Detection Theory to the analysis of recognition memory data; this is a statistical procedure adopted earlier in psychophysics (Green & Swets, 1966) for separating sensitivity (memory strength) from bias. Mandler also had long-standing interests in emotion and in the history and philosophy of psychology as a science, and eventually he was one of the first cognitive psychologists to bring CONSCIOUSNESS within the scope of inquiry. In 1966 Mandler recruited Donald Norman from Harvard and hired a new Ph.D. from the University of Toronto, Peter Lindsay. They were soon joined by David Rumelhart, a mathematical psychologist who worked with Estes, receiving his Ph.D. from Stanford in 1967. These three created the LNR (Lindsay–Norman–Rumelhart, or ELINOR) research group, which included as graduate students and outside visitors several researchers who later played important roles in cognitive science. The group's 1975 book, *Explorations in Cognition*, ended with a suggestion that the "concerted efforts of a number of people from ... linguistics, artificial intelligence, and psychology may be creating a new field: *cognitive science*" (Norman and Rumelhart, 1975, p. 409). This prescient statement, plus the subtitle of a 1975 book edited by Daniel Bobrow and Allan Collins, are the earliest published uses of the term *cognitive science* that we have identified.

The LNR group pursued research across a broad sweep of topics and methodologies. There were mathematical models of word recognition, mathematical and computational models of analogy, and studies of memory, perception, imagery, sentence processing, story understanding, and more. One of the main joint concerns of the group was the organization of information in long-term memory, and they developed a semantic network format for representing information that was a descendant of Ross Quillian's work in artificial intelligence (section 2.3.3). In a typical semantic network, nodes representing concepts are connected by labeled, directed relations. To this the LNR group added procedural information specifying how tasks should be performed (in the same format as was used for knowledge representation) and a notation for schemata such as those inspired by Bartlett (see above) and Piaget, yielding what they called *active structural networks*. The active semantic network was not just a theoretical proposal: a version named MEMOD was partially implemented on a computer: "This was a major project in itself, and the resulting program became one of the largest and most complex interactive programs operating on the campus computer center computer (a Burroughs 6700)" (p. xiv).

The 1975 book reporting this work juxtaposes primarily experimental chapters with primarily theoretical chapters in several areas. Topics included the active structural network, MEMORY, PERCEPTION, REPRESENTATIONS of knowledge, PROBLEM SOLVING, LINGUISTIC THEORY, LANGUAGE PROCESSING, and augmented transition network parsers. In the middle of the book were three chapters from the "verb group," which had developed semantic representations for verbs composed of primitive predicates that were linked to their arguments via labeled relations inspired by Charles Fillmore's case grammar (e.g., agent, object, source, goal). Rumelhart and James Levin used these representations in developing a computer-implemented language comprehension system. Dedre Gentner focused on fractionating verbs of possession into their primitive predicates, and demonstrated that children's learning of those verbs progressed from those involving fewer components (e.g., *give* involves an agent **doing** something that **causes** a **transfer** of goods) to those including more (e.g., *sell* adds a **contract** for **transfer** of money). The relatively simple representation of *give* is illustrated in figure I.8. Adele Abrahamsen proposed a method for constraining the analysis of verbs of motion into component predicates: she asked subjects to recall a story containing these verbs and used recall errors to suggest components that had been added or deleted from the original verbs.

Other research groups were also part of the cognitive science community at UCSD. The psychology department itself had several strong experimental and social psychologists. At the nearby Salk Institute, Ursula Bellugi, also a product of the Harvard Center, began psycholinguistic studies of American Sign Language in the early 1970s. With such collaborators as Edward Klima and Susan Fischer (MIT-educated linguists) and Patricia Siple (a recent UCSD Ph.D. in psychology), she did much of the pioneering work showing that American Sign Language was a true language both in structure and processing. (Later this topic was addressed in the psychology department by Elissa Newport and Ted Supalla, in the medical school by Helen Neville, and at Salk by Howard Poizner.) In the linguistics department, there was strong influence from Chomsky and also ongoing discussion of *generative semantics*. Finally, Roy D'Andrade was a pioneer in COGNITIVE ANTHROPOLOGY, and Aaron Cicourel pursued a cognitive approach in sociology (see Article 30, ETHNOMETHODOLOGY).

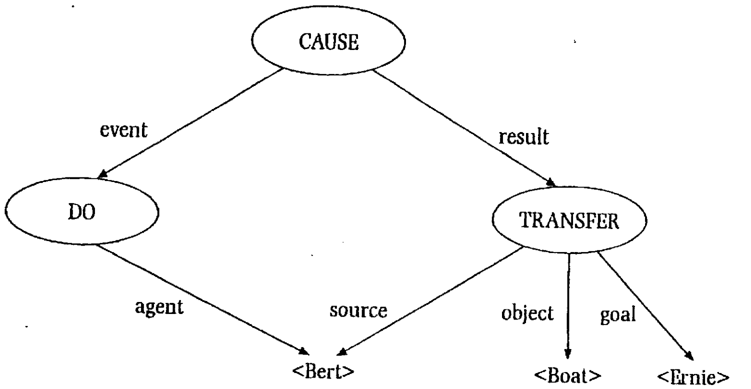

Figure I.8 The LNR verb group's representation of the meaning of the verb in the sentence "Bert gives a boat to Ernie." Each entity (angle-bracketed nodes) has a labeled case relation to a primitive predicate (oval nodes).

The creation of an Institute for Cognitive Science at UCSD was not to occur until the late 1970s (becoming a department in 1987), but the psychology graduate program had been designed from its beginning in the mid-1960s to encourage collaborative work with faculty, interaction across disciplinary boundaries, and even interaction across institutions: the LNR group had fruitful contacts with researchers on both coasts during the exciting period of the early 1970s. On the East Coast, this included cognitive scientists at Bolt Beranek and Newman outside Boston (especially Allan Collins, William Woods, Joseph Becker, and the late Jaime Carbonell), Ronald Kaplan at Harvard University, and more sporadic contacts with the very active AI lab at MIT. On the West Coast, there were three working visits with Stanford's cognitive psychologists and also fruitful interactions with Stanford-area AI researchers (fostered especially by Norman, who coauthored publications with Bobrow and with Schank).

2.4.3 *University of Minnesota*

The University of Minnesota provides an interesting alternative model of the development of cognitive psychology. It was a department with strong roots in learning theory and long-time home of two major centers: the Center for Research in Human Learning and the Institute for Child Development. Not as strongly identified with cognitive science as Stanford or UCSD, it nonetheless generated some important research in this field.

One of the leaders of the move towards cognitivism at Minnesota was James J. Jenkins, who began his career by trying to develop a mediational theory of learning in the behaviorist tradition that would be sufficient to explain the acquisition of language. The mediators were envisioned as internal stimuli and responses, a series of which was thought to facilitate the ability to make associations needed for language. But Jenkins discovered that more complex mediations failed to develop automatically, in the manner necessary to explain language, and as a result began to abandon behaviorism. He also began to recognize that the sort of grammar for which mediation theory was designed, a slot grammar of the form advocated in structural linguistics, was being superseded by Chomskian transformational grammars, which required rules, not associations. Jenkins spent the 1959–60 academic year at the Center for Advanced Studies in the Behavioral Sciences (as did George Mandler), the same year in which Miller, Galanter, and Pribram were writing *Plans and the Structure of Behavior*. Then in 1964–5 Jenkins spent another year at the Center, this time in the company of Jerry Fodor and Sol Saporta. Fodor (section 2.6) was a philosopher who had by this time become a major expositor of Chomsky. At the Center he and Saporta directed a tutorial reading of *Syntactic Structures*.

Returning to Minnesota, Jenkins was appointed director of the new Center for Research in Human Learning. One of the first activities of the Center was a summer program which involved a number of recent Minnesota Ph.D.s. The outside instructors were Fodor, Walter Reitman, and David Premack. Reitman's role was to introduce new developments in artificial intelligence, using his then new textbook *Cognition and Thought*. Fodor advanced the new cognitive views, especially those of Chomsky, while Premack defended somewhat more traditional behaviorist positions. Fodor and Premack sparred throughout the summer. Walter Weimer, then a graduate student at Minnesota, describes the interactions this way:

> Premack represented one direction – trying to resuscitate behaviorism and be "sophisticated" by saying we can address all those higher mental processes, too. Premack was playing catch-up ball, but without knowing he was playing catch-up against Dallas in the last minutes of the Super Bowl. Another direction was Fodor, who just stole the show. He was already so involved in the linguistic argument that he was doing normal-science transformational work, that hot-off-the-mimeo-presses, "This is the latest grammar from MIT." He was already doing normal science in the new linguistics, and saying to us "All right, this is what you must do in psychology to catch up." (Interview in Baars, 1986, p. 300)

Jenkins has directed more than 70 dissertations, in almost every area of cognitive psychology. Two of them were by John Bransford and Jeffrey Franks, who carried out a series of influential experiments that provided evidence for internal cognitive processes. Drawing upon Michael Posner and Steven Keele's demonstration that subjects could abstract a prototype they had never seen from random dot patterns that were generated from the prototype, Bransford and Franks (1971) developed their own stimuli – geometrical patterns that were derived from a prototype in a rule-governed way – and showed that when subjects were asked to draw the typical figure they had just seen, they drew the prototype they had not seen. From this starting point in visual perception, Bransford and Franks moved to language. They created complex sentences such as "The ants in the kitchen ate the sweet jelly which was on the table" to function as prototypes. They then extracted the four component propositions (e.g., "The jelly was sweet") and recombined these into derived sentences expressing one, two, or three of the propositions. Subjects were presented with subsets of the derived sentences from more than one prototype sentence in scrambled order. Then they were given old and new sentences to recognize, in which the new sentences were combinations of the four component propositions that had not yet been presented. Not only were subjects poor at distinguishing between old and new sentences; they actually gave their highest confidence ratings to the never presented prototypes. In a later study, Bransford, Barclay, and Franks (1972) showed that never presented inferences also were prone to false recognition. These studies were taken as evidence that memory is constructive.

2.5 *Learning to think: artificial intelligence (AI)*

At the time of the 1956 Dartmouth Conference, the Newell–Simon–Shaw Logic Theorist was the only functioning AI program. The 1960s and 1970s, though, saw rapid expansion. During this period, three centers for AI research assumed prominence: Allen Newell joined Simon at Carnegie–Mellon in the Graduate School of Industrial Administration; Minsky and McCarthy joined forces to create the Artificial Intelligence Group at MIT; and McCarthy left MIT in 1962 to join the computer science department at Stanford and founded an AI laboratory there. Artificial intelligence research during this time was expensive, since running the programs was computationally very demanding, and the existing computers were considerably less powerful than contemporary notebook computers. One of the major sources of funding was the Advanced Research Projects Agency (ARPA) of the Department of Defense, created after Sputnik in an attempt to ensure US competitiveness in science and technology. Unlike other granting agencies, ARPA did not rely on peer review and often focused on funding individuals thought to be promising rather than specific projects. One of its first major grants was a $2,220,000 grant to MIT for a project known as MAC (for both Multiple Access Computer and Machine-Aided Cognition). A third of this money went to the Artificial Intelligence Group; given McCarthy's departure, this money went to support Minsky and his new collaborator, Seymour Papert. We will return to Minsky's work below, but first we need to capture some of the atmosphere of early AI.

As soon as AI began to develop, it attracted attention in the popular press. One of the most visible domains into which AI ventured was game playing. Checkers and chess, in particular, are games of strategy that seem to require intelligence; thus, if AI systems could succeed in these games, this might be taken as evidence that artificial intelligence is possible. One of the early leaders in these efforts was Arthur Samuels, who in 1946 left Bell Laboratories for the University of Illinois. Aware that computers were on the horizon, he sought to have the university either buy or build one for him. As part of a scheme to raise money, his research group came up with the idea of building a computer to play checkers and then, by defeating the world checkers champion, raise additional money. Although they failed to develop either the hardware or the program, Samuels was bitten by the bug to build a checkers-playing computer and continued pursuing the idea when he went to IBM in 1949. When in 1951 IBM produced the 701, its first commercial computer, one of the programs prepared to run on it was a crude checkers-playing program by Samuels. Samuels kept improving on the program, partly by providing it with a capacity to learn from previous games, and by 1961 it was playing at master's level.

As challenging as checkers turned out to be, it was chess that emerged as the holy grail for AI. Early on it was realized that chess, unlike checkers, could not be played successfully by simply considering all possible sequences of moves that constituted games and choosing only sequences that led to winning (or at least not losing) – for the simple reason that there are too many such sequences (on the order of 10^{120}). In principle, one could represent all possible games by constructing a tree structure in which chess positions are nodes and each possible move from a position is a branch from its node. Proceeding from the root of the tree (initial chessboard position), a path down a series of branches to a terminal node would represent a single game. To make search of such a tree manageable, a procedure was required to prune it – that is, to rule out certain branches, and all branches that branch off from them, as unhelpful and thus not meriting further consideration. In a *Scientific American* article in 1950 Claude Shannon had proposed a strategy of following out all paths to a certain depth in the tree (e.g., four moves), evaluate the resulting board positions, and then use a minimax strategy to choose the best next move. (A minimax procedure is one which selects the move whose worst outcome is better than the worst outcome of all competitors.) Variants of Shannon's strategy permitted continued search of those branches which either seemed to involve potential risk or held the most promise. Other AI researchers quickly joined in the effort to develop a championship-level chess program; for example, even before they had developed Logic Theorist, Newell, Simon, and Shaw had embarked on developing a chess-playing computer. In 1957 they predicted that a computer would be world chess champion (if the rules allowed it) within a decade. But by 1967, chess-playing programs were still so inadequate that international master David Levy confidently bet that no computer would beat him within a decade. He not only won that bet, but won a renewed bet. Finally (shortly after the penultimate draft of this introduction was written), an IBM system dubbed "Big Blue" won a rematch with world champion Garry Kasparov on May 11, 1997 – four decades after the original optimistic prediction. Ironically, it appears that Big Blue's victory was clinched by its ability to unnerve Kasparov, rather than to outthink him; the opponents were tied when Kasparov fell apart in the sixth and last game.

2.5.1 *Simulating human performance*

Despite the crowd-pleasing appeal of such contests, game-playing programs have lived at the periphery of AI. Far more central was the goal of simulating human REASONING in a variety of domains. Throughout the 1960s and beyond, Newell and Simon at Carnegie-Mellon were a major presence at the intellectual center of this effort. As a strategy for developing computer programs that simulated human performance, they adopted a strategy developed by two psychologists, O. K. Moore and S. B. Anderson (1954), of asking subjects to *think aloud* while solving puzzles (see Article 33, PROTOCOL ANALYSIS). From the resulting reports they extracted reasoning strategies or heuristics to be incorporated into their next program, General Problem Solver (GPS). As its name suggests, GPS was designed to employ general procedures that could be widely applied in PROBLEM SOLVING. GPS was designed as a PRODUCTION SYSTEM, which consists of a working memory and a set of production rules. The working memory provides for the temporary storage of symbolic expressions. The heuristics were programmed in the form of production rules that paired a condition with an action: "If circumstances X obtain, do Y." X is an expression that might appear in working memory. If it did, then the production rule would *fire*, directing the system to perform action Y, which might involve adding or removing content from working memory or sending output from the system.

One kind of problem to which Newell and Simon applied GPS was cryptarithmetic. The challenge in cryptarithmetic is to replace the letters with numbers in an equation such as

$$\begin{array}{rl} & DONALD \\ + & \underline{GERALD} \quad D = 5, \\ & ROBERT \end{array}$$

so that the result is a valid addition problem. The key to their approach was *means–end* reasoning wherein the machine compares a description of the goal state with a description of its current state; if there is a difference, it employs a variety of operators to reduce this difference. One important aspect of the strategy involves working backwards: if no operator will take the system directly to the goal state, then operators are identified which would take it to the goal state from another state. Then a goal is established of reaching this new state, and again operators are sought to reduce the difference. Once GPS was up and running, Newell and Simon devoted much of their effort to comparing its performance to that of humans and revising the program to better simulate human performance. This project culminated in the 1972 publication of *Human Problem Solving*.

GPS was grounded on the assumption that intelligent behavior stemmed largely from general reasoning principles, not detailed knowledge of particular circumstances. Increasingly during the 1970s, though, AI researchers began to recognize the relevance of specific knowledge in solving problems. One direction of this effort led to the development of expert systems, programs that would incorporate knowledge gained from interviewing human experts in a particular domain. The designers would represent the expert knowledge in the form of rules and then apply general reasoning principles to arrive at inferences and answers to queries. Often a significant amount of additional manipulation of the system was required to get it to perform well, but some of the systems ultimately equaled or exceeded the performance of human experts. An exemplar of such research is DENDRAL, developed at Stanford University and described in 1971 by three researchers: Simon's student Edward Feigenbaum, philosopher-turned-computer scientist Bruce Buchanan, and Nobel Laureate geneticist Joshua Lederberg. DENDRAL performed at an expert level in analyzing data from mass spectrographs to determine the molecular structure of the organic compound being analyzed. In another project that relied on expert knowledge, Buchanan's graduate student Edward Shortliffe (1976) developed MYCIN, a program for diagnosis of infectious blood diseases.

Many of the early pioneering efforts in AI occurred at MIT during the period when McCarthy and Minsky were both there. John McCarthy developed what was to become the standard language for AI programming, LISP (LISt Processing language). It took from Newell and Simon's language IPL the idea of working with lists, each item of which would index the next item. LISP advanced beyond IPL in part by utilizing the lambda calculus, which allows functions to be treated as objects and hence as arguments in other functions. Another advance was to allow the items to represent not just simple imperatives (do this) but conditional statements like those in production systems (if specified conditions are met, do this).

In 1968 Minsky edited an influential book, *Semantic Information Processing*, in which McCarthy gave an overview of LISP, and a number of doctoral students published abbreviated versions of their dissertations. In general, each dissertation involved a computer program written to simulate some aspect of human cognition. For example, Thomas Evans's program ANALOGY was designed to solve visual geometric analogy problems in which it had to pick one of five possible solutions to problems of the form "*A is to B as C is to?*" Descriptions of the geometrical forms were represented in propositional expressions in LISP. The program then identified the difference between the representations of A and B and applied this difference operation to the representation of C to arrive at its answer.

While adequate programming languages were important to the development of AI, it was also necessary, if AI systems were to appear intelligent to humans, that users be able to interact with them in ordinary language. Daniel Bobrow's MIT dissertation, also included in the 1968 book, described a program called STUDENT which solved algebra story problems such as: "Bill's father's uncle is twice as old as Bill's father. Two years from now Bill's father will be three times as old as Bill. The sum of their ages is 92. Find Bill's age." In order to solve the problem, STUDENT needed first to transform the story problems into equations, which it did by matching the sentences against stored templates (e.g., ____ times ____) and from that extracting the equations.

Programs like STUDENT had to *kluge* (a term of art popularized at MIT for ad hoc computational solutions) serious problems, such as really understanding English. Matching sentences against stored templates would work sometimes, but sentences were easily constructed on which the strategy failed. To move from these *toy* cognitive activities to cognitive activities on the scale of real life required providing AI systems with vision and a more principled ability to understand natural language. Some early AI researchers thought vision, at least, would be easy, since even simple organisms could sense features of their environment, and modestly more complex organisms could detect visual layouts and recognize objects. Simplicity in organism was thought to betoken simplicity in task. Minsky, who had received ARPA funding for vision research, therefore assigned it as a summer project in 1966 to a precocious undergraduate, Gerald Sussman.

Soon realizing the challenge of the task, Minsky and Papert made the development of an AI system for vision a major part of their mission. To make the project more tractable, they developed the idea of using a simplified visual world as the target: the blocks micro world. Blocks were chosen because of their straight edges and relatively smooth appearance. The blocks world also became a target for work in robotics (getting robots to move blocks around) and natural language processing (with the goal of eventually using English to direct robots' interactions with the blocks). One successful program, written by Patrick Winston for his dissertation, used semantic networks of the sort developed by Quillian (see below) to acquire information about the blocks world linguistically. For example, Winston's program would be presented with a symbolic description of a configuration such as an arch and would be told that it was an arch. It would also be presented with some non-arches. After being fed several examples, the program was able to form a general description that would include the arches and exclude the non-arches.

Developing AI systems capable of navigating real environments was the objective of a major AI research group located at the Stanford Research Institute (which subsequently dropped its official affiliation with Stanford and was renamed SRI International). Also heavily supported by ARPA funding, Charles Rosen headed a team that included Bertram Raphael (who received his Ph.D. with Minsky), Richard Fikes, and Nils Nilsson. In 1969 they built a robot, Shakey, which propelled itself on wheels and used a TV camera for visual input. Although mobile, Shakey was restricted to a suite of seven rooms, some of which contained boxes that Shakey could push around or stack. Fikes and Nilsson developed a control system for Shakey called STRIPS (Stanford Research Institute Problem Solver), which consisted of three-part rules: one part would specify preconditions (*x* is on the table, nothing else is on *x*, and Shakey's hand is empty), a second part would delete conditions (*x* is on the table and Shakey's hand is empty), and a third part would add conditions (Shakey is holding *x*). Goals would also be specified in simple predicate structures, and a goal would invoke a rule if the rule contained as one of its add conditions a state specified in a goal. STRIPS could reason backwards from goals to preconditions, set the preconditions as goals, etc., until it reached preconditions that could be executed. Further, once a plan was formed in STRIPS for reaching a goal, that plan could be stored, with variables replacing names, so as to be employed in similar circumstances in the future. STRIPS thus contained many of the features of **PRODUCTION SYSTEMS**.

Perhaps the most impressive research with a blocks world, however, used a simulated rather than real blocks world and robot in order to focus on how a program could communicate in English about acting in a blocks world. For his dissertation under Papert, Terry Winograd (1972) used a computational representation of a blocks world which his program, SHRDLU, could also display on a computer monitor. SHRDLU would process sentences supplied to it by applying both syntactic and semantic rules of unprecedented sophistication. SHRDLU was able to answer a broad range of questions about the blocks world and could also carry out actions in response to commands. Some of the commands could be satisfied by a single action such as grasping a block; others required a sequence of preliminary actions. An important aspect of SHRDLU was its underlying mode of operation. It did not employ the theorem-proving approach of such investigators as Newell, Simon, and McCarthy. Instead, using PLANNER, a LISP-based language, it operated with various subprograms, each of which pursued an independent goal; a subprogram could, in appropriate circumstances, take control of the operation of the program until its goal was realized or failed.

2.5.2 *AI aims to get real*

With programs such as GPS, STUDENT, and SHRDLU, AI had clearly achieved some success. But the hype for early AI had been much greater: the promise of a computer becoming world chess champion in the 1960s and of demonstrating humanlike intelligence by the 1970s. When these aspirations were not realized, funding reductions and critical assessment began. In the mid-1970s ARPA reduced its funding of AI research. But perhaps as important was critical assessment within and outside the AI community. One of the major problems facing AI was identified by John McCarthy in collaboration with philosopher Patrick Hayes (McCarthy and Hayes, 1969). McCarthy's approach to AI – using formal logical derivations to arrive at actions – required a complete internal representation of all relevant features of the world, a frame of reference. Difficulties ensued because activities in the world, including those initiated by the AI system, could change this frame. McCarthy and Hayes labeled the problem of how to update the frame of reference the *frame problem*. Solving it seemed to require providing the computer with complete knowledge about what would change and what would remain the same in the world as a result of an activity. Providing such knowledge would be, at best, a mammoth undertaking. What the frame problem indicated was that successes like SHRDLU depended critically on the limitations imposed on the micro-world in which it operated. Since the behavior of that world was itself controlled by deterministic laws in the computer, it was possible for the system to update itself on all changes. But this would not carry over when an AI system was operating in the real world – as the brain of a robot, for example.

Almost from its inception, critics of AI emerged who questioned whether problems like the frame problem could ever be solved. The most influential of these was the philosopher Hubert Dreyfus. Dreyfus had an early clash with Herbert Simon after a colloquium at MIT in 1961, where Dreyfus was teaching. Dreyfus was then hired as a consultant by the RAND Corporation for the summer of 1964 and produced a report entitled "Alchemy and AI," which, as its title indicates, challenged the legitimacy of the whole AI enterprise. (Seymour Papert wrote a response, entitled "The artificial intelligence of Hubert L. Dreyfus: a budget of fallacies," which RAND decided not to publish for fear of a libel suit, but which later appeared as an MIT technical report.) Dreyfus's report became the basis for his 1972 book *What Computers Can't Do.*

Dreyfus's objection went to the very foundation of AI. He questioned whether "processing data representing facts about the world using logical operations" was sufficient to account for what our cognitive systems do. Part of Dreyfus's critique involved noting that for an AI system working on logical operations, no meaning is attached to the data structures or symbolic representations. To deal with meaning, he contended, one had to get beyond formal rules and deal with the world and body as experienced – what philosophers call *phenomenological experience*. One of the features of our phenomenological experience is what Dreyfus (in the manner of William James) calls *fringe consciousness*, our awareness of features of the world that are not focal (i.e., features to which we are not explicitly attending but which nonetheless influence our focal consciousness). He contended that the effects of fringe consciousness could not be simulated by simply providing more symbolic representations and rules; to have fringe consciousness, one must have a body operating in a world. A long-distance runner, for example, may respond to the twists and turns of the road on which she runs without consciously representing those twists and turns to herself. For humans, cues and constraints from the world shape our understanding, and these are unavailable to the AI system – or so claimed Dreyfus.

While the objections raised by Dreyfus sparked bitter controversy, some of his concerns began to be addressed in research during the 1970s. One of the limitations of early AI programs was that they essentially treated each piece of information as a separate proposition, which would then be related to other propositions through associations or logical relations. One remedy was to develop more complex data structures that represented information in relation to other information. An idea of how to develop relational data structures – Ross Quillian's semantic network approach – was actually developed quite early. Quillian developed his basic framework as a graduate student under Herbert Simon in the mid-1960s, and a modified version of his dissertation was published in Minsky's (1968) book *Semantic Information Processing*. Quillian thought in terms of a network of nodes. For each sense of a particular word (a *word concept*) there was one type node (similar to a dictionary entry) and numerous token nodes (for each particular use of the word concept). The node for one word concept was connected to nodes for other word concepts that figured in its definitions, and these to nodes that figured in their definitions, creating a configuration of links. One method used by Quillian to explore the behavior of such networks was to write a computer program to compare any two word concepts. Starting at the two type nodes, the program worked its way outwards and placed activation tags on each node encountered. The overlaps in meaning between the two word concepts were determined in this way, and Quillian even wrote another program to give a crudely worded English report of the comparison. He also wrote a program that did a fair job of disambiguating words in sentences. Figure I.9 provides a relatively simple example of what happens when two related concepts are compared: the two paths traversed by moving out from *cry* and *comfort* converge at *sad.*

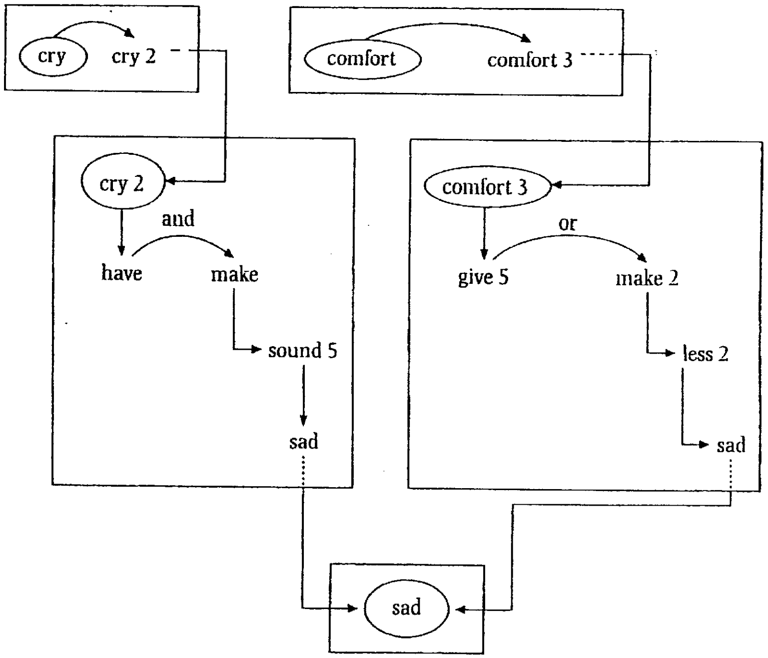

Figure I.9 A portion of a semantic network as explored by Quillian (1968). Two concepts, *cry* and *comfort*, are shown to be related in that paths lead from both of them to *sad*.

Quillian's networks were even more computationally demanding than other early AI programs; given the limitations of the available computers, he never got beyond the demonstration phase for this particular network design. But his ideas had enormous influence. We have already encountered two later projects that used modified network designs to simulate human performance: the LNR group's ELINOR model at UCSD and Anderson and Bower's SAM model at Stanford. Additionally, Quillian collaborated with psychologist Allan Collins to predict human reaction time patterns from a simplified version of Quillian's network (Collins and Quillian, 1969). This spawned the spreading activation research paradigm, which has had a long, vigorous life within mainstream experimental psychology.

One of the prime movers in developing larger-scale knowledge structures was Roger Schank. Originally trained at the University of Texas as a linguist, he broke fundamentally with the Chomskian focus on autonomy of syntax and reinvented himself as an AI researcher. Radically de-emphasizing syntax, Schank argued that meaning representations could be extracted directly from sentences. As an assistant professor in linguistics and computer science at Stanford, Schank tried to write computer programs capable of understanding natural language. Analyzing the meaning of verbs, he proposed eleven primitive actions that could be used to obtain semantic representations for a large number of verbs. Examples included PTRANS (transferring the physical location of an object), MTRANS (transferring mental information within or between subjects), and ATRANS (transferring the abstract location of an object; i.e., possession). Each primitive action involved roles such as actor, object, source, and goal. When the program needed to interpret a particular verb in a sentence, it started constructing a meaning representation containing the appropriate primitive actions; the specified roles for those actions would then need to be filled. If the program could not fill them, it would reassess the actions that had been assigned to the verb. For example, *sold* frequently involves an ATRANS, but in the phrase *sold out* it does not. Schank therefore attached other general rules to the various predicates, which enabled the system to go beyond the literal meaning of a sentence to make plausible inferences.

Schank's first effort to employ such rules in sentence interpretation, MARGIE (Memory, Analysis, Response Generalization in English), suffered from the fact that typically there were too many plausible inferences that might be licensed. After moving to Yale, Schank began a fruitful collaboration with psychologist Robert Abelson. They overcame these limitations by developing higher-order knowledge structures, *scripts*, which characterized the general structure of events in a common experience from the point of view of a specific participant (Schank and Abelson, 1977). Their best-known script was for a diner going to a restaurant. Since the events that transpire in a restaurant depend upon the kind of restaurant, the restaurant script contained tracks for fast food restaurant, cafeteria, coffee shop, etc. Within each track were a number of roles and a number of scenes (e.g., entering, ordering, eating, and exiting), as well as a typical sequence of primitive actions. The following is the portion of the coffee shop track of the restaurant script in which the customer (S) requests a menu from the waiter (W) and the waiter begins to respond:

S MTRANS signal to W
W PTRANS W to table
S MTRANS "need menu" to W
W PTRANS W to menu

Although only a few of the primitive actions might be mentioned specifically in a story, Schank and Abelson proposed that readers supply the others by using scripts to understand the story. Hence, readers should be able to answer questions about primitive actions not mentioned in a story and include some of these in a paraphrase of the story. (See Article 51, REPRESENTATIONS.) To deal with nonstereotyped circumstances, Schank and Abelson added the recognition that humans generally have goals, as well as plans for meeting them.

Schank's graduate students proceeded to develop a number of programs that used scripts to understand stories, including Gerry DeJong's FRUMP, which analyzed stories on the UPI news wire, and Janet Kolodner's CYRUS, which emulated government official Cyrus Vance. Employing FRUMP to learn from the UPI about Vance, CYRUS could then answer questions and make inferences about Vance. (Subsequently Schank's research led him to recognize the importance of even larger structures encoding knowledge of previous experiences of a kind, such as previous meals one has cooked. These cases provide the basis for CASE-BASED REASONING models. In addition, since moving to Northwestern in 1989 to direct the Institute for the Learning Sciences, Schank has become increasingly concerned with applications of AI to EDUCATION.)

Another proposal for higher-level knowledge structures was found in Minsky's (1975) work on frames and frame systems. A frame is a data structure representing stereotypical situations in terms of features that are always true of such situations, as well as terminal slots for features which may take on a variety of different values but must be assigned some value in a given situation. For example, in looking at a room from a given viewpoint, there will always be walls in the scene, but several options for the color of the walls. Slots will generally have default values associated with them, but these can be dislodged if alternative information is presented. The slots themselves may be filled by frames, providing for a recursive representational system. Another source of systematicity is that frames are related to one another by transformations (e.g., the transformation of moving from one viewpoint in a room to another). Some transformations will result in changes in the slot-fillers, while others will retain the same values. Minsky proposed that when a person encounters a situation, what the person tries to do is to match the information about the situation to a frame in a stored frame system; once a possible frame is proposed, it generates expectations that guide further search for information. If such a search produces information inconsistent with the frame, then a new proposal for a frame must be advanced. Minsky linked his notion of a frame to *schemata*, which figured in Bartlett's (1932) account of memory, and to *paradigms*, which figured in Thomas Kuhn's (1962/1970) account of normal science.

The introduction of large-scale knowledge structures such as semantic networks, scripts, and frames went part way toward addressing Dreyfus's concerns, as he acknowledged in an extended introduction to the second edition of his book *What Computers Can't Do* in 1979. But they still did not meet his objective of having systems engage the real world through real bodies. Without that, Dreyfus contended that such systems would never achieve real knowledge. An extreme example of this can be found in the program ELIZA, designed by Joseph Weizenbaum to carry on conversations with human interlocutors without any understanding of the topic of the conversation. Weizenbaum pulled this off by modeling ELIZA after Rogerian psychotherapists, who are known for their technique of nondirectively reflecting back to the client what the client has said. One strategy was to insert "Why do you think ____" before repeating what the interlocutor said. Another would monitor for key words such as *mother* in the interlocutor's statements and respond "Tell me about your ____." The program would also make appropriate substitutions, such as *your* for *my*, so as to mimic a real conversation.

Many people who interacted with ELIZA were seduced into thinking that it understood what they were saying and would engage in elaborate conversations. In those cases, ELIZA seemed to be passing the Turing test (interacting indistinguishably from a human), and this was impressing people far more than Weizenbaum thought it should. In *Computer Power and Human Reason* (1976) he strongly criticized the tendency in AI to overinterpret performances achieved using symbolic representations and rules. Because AI systems lacked any understanding of the symbols they were manipulating, they might imitate humans (to a degree) but could never really replicate human intelligence. The moral risk, then, was to be taken in by mimicry and turn over to machines decisions that required true human intelligence.

2.6 *Getting a philosophy*

Philosophers have theorized about the mind for 2,500 years; thus, it may be a bit surprising that we have reached this point in describing the emergence of cognitive science and have yet to focus on the contributions of philosophers. One major reason for this is that the endeavors that motivated the development of modern cognitive science were mostly empirical efforts: experiments and other data-based studies in psychology and neuroscience, analysis of sentences judged to be well formed in linguistics, and construction of machines (or programs for machines) that could carry out cognitive activities in AI. Thus, it might seem that cognitive science, like so many other sciences, would leave philosophy behind after developing its empirical side. But this was not to be: during this period of initial maturation of cognitive science, philosophical inquiry (sometimes by nonphilosophers) came to play an increasingly important role.

Perhaps the oldest philosophical problem relevant to cognitive science is the mind–body problem, which took its modern form with the seventeenth-century philosopher René Descartes' contention that the mind is distinct from the body. Descartes thus defended a dual substance ontology, which raised as a central question how nonmaterial minds could interact with physical bodies, including brains. The primary opposition to dualism has come from theorists labeled collectively as *materialists*; materialism, however, is such a broad camp that it includes a number of very different views about how mental states relate to physical ones, including how psychological states relate to states of the brain. One very prominent position, advanced by Gilbert Ryle in the 1940s, seemed to comport well with the behaviorism then current in psychology. Ryle proposed that mental predicates (such as *believes*) do not designate internal states of agents but rather describe propensities of agents to behave in certain ways. This approach, however, seemed incapable of accounting for other mental states, such as sensations. To handle sensations, in the 1950s philosopher J. C. Smart and philosophically minded psychologist U. T. Place advanced the mind–brain *identity theory*, which identified sensations with brain states; for example, pain was simply a particular brain state. The identity theory eventually became generalized to other mental states such as beliefs. At the time of the birth of cognitive science, the identity theory was the dominant version of materialism advanced by philosophers.

But the identity theory did not seem to fit well with a major feature of the new cognitive approach: if mental states were identical with brain states, then particular mental states were limited to organisms with similar brains. Many cognitivists were attracted to the idea that the same sort of cognitive state or condition could occur in diverse brains or central nervous systems. Different sorts of animals may perceive the rainfall, taste a lemon, attend to a red triangle. Those attracted to AI were further enticed by the fact that computers too might perceive, attend, and reason. Reflecting on animal minds and the possibility of artificial minds, as well as minds in other organisms, Hilary Putnam, a philosopher who began his career at Princeton and then moved to

Harvard, argued against the identity of mental states with brain states (Putnam, 1960). He claimed that states such as pain and hunger could occur in organisms whose brains are as different from ours as those of octopuses. He thus introduced the idea that mental states might be *multiply realizable*, and consequently cannot be identified with any of their realizations.

Rejecting the identity between psychological states and brain states might seem to be part of an argument for dualism, but Putnam certainly did not construe himself as abandoning materialism. What, then, are mental states if they are not to be identified with brain states? Putnam (1967) proposed that mental states be identified in terms of their causal or functional role in mediating between sensations and behavior. That role – rather than its neural realization – determines the nature or identity of the state. For example, what determines whether a particular cognitive process or state is a preference for an ice-cold beer, as opposed to the perception that a flower is red, is that the state is caused by thirst and gives rise to other mental states such as planning to move towards the refrigerator or requesting that someone bring over a beer. The state causes other mental states and then beer-seeking behavior, because, roughly, it is a preference for beer. The brain must somehow implement the causal role, but it is the role – not the brain – that makes the state a beer preference.

Putnam dubbed his theory *functionalism*. His use of the term is confusing in some respects. Within psychology, for example, we have seen that *functionalism* is used to describe an approach to psychological theory construction stemming from James that has an evolutionary orientation. What Putnam had in mind was not Darwin's work on adaptivity, but rather Turing's work on machine tables. However, the essence of Putnamian functionalism is strikingly clear, and the term aptly captures that essence. For Putnam cognitive states are defined by their typical causes and effects. (Other defenders of functionalism, such as Daniel Dennett and William Lycan, maintain Putnam's emphasis on causal role, but, as a result of allowing for multiple iterations of functional decomposition, allow for a greater connection between psychological functional roles and the neural processes that realize them.)

Functionalism fits well into the cognitive agenda. By drawing links between mental processes and operations in computers, functionalists saw that one could avoid the chauvinism of limiting cognitive states to systems with our kind of brain while seemingly also avoiding some of the problems of behaviorism (see, however, Block, 1978, for an influential criticism that denies the potential for finding such a middle ground). Functionalism (in one form or another) was rapidly endorsed by many philosophers, who acclaimed it as providing the cognitivist solution to the mind–body problem. But in the hands of one of Putnam's students, Jerry Fodor, it became just one of several contentious claims in an ambitious theory that attempted to answer a key question for cognitive science: "What would a satisfactory theory of cognition look like?" Neisser and others had asked the question and sketched suggestive answers. In *The Language of Thought* (1975) Fodor, then at MIT, offered one of the most detailed answers to that question, one that has continued to play an influential, often divisive role in cognitive science. The heart of his answer was something he called the *Language of Thought* (LOT) hypothesis. Here is a capsule characterization:

> LOT: To be a cognizer is to possess a system of syntactically structured symbols-in-the-head (mind/brain) which undergo processing that is sensitive to that structure. Cognition,

> in all of its forms, from the simplest perception of a physical stimulus to the most complex judgment concerning the grammaticality of an utterance, consists of manipulating symbols-in-the-head in accord with that syntax. The system of primitive, innate symbols-in-the-head and their syntactic combination in sentence-like structures is sometimes called "mentalese."

Fodor sometimes referred to the syntactically structured symbols-in-the-head as *representations* (see Article 50, REPRESENTATION AND COMPUTATION). By this he meant that these symbols represent the world by referring to things and by predicating or ascribing properties to them. "Water is wet" (or its mentalese equivalent), for example, refers to water and attributes the property of being wet to water. This is the meaning of the sequence of symbols (its semantics). An important claim, however, is that semantics need not be accessed when operating on the symbols: operations are licensed solely on the basis of form (the "shape" of each symbol and the syntax by which they are combined).

Since 1975 Fodor has put forward a number of arguments for LOT. Initially, many of his arguments focused on learning. He argued that in order for a cognitive system to learn a language like English, for example, it had to be able to advance hypotheses (e.g., *water* means water) and then test these. Thus, the cognitive system needed a mode of representation capable of expressing any hypothesis it might want to test. One consequence of this kind of reasoning was that LOT could not be a natural language like English, but had to be available to the system before it could acquire language. Fodor therefore followed another of his mentors, Chomsky, in arguing for a strong nativism. While Chomsky had argued that knowledge of basic grammatical rules must be innate, Fodor argued that a powerful representational system had to be innate (see Article 45, INNATE KNOWLEDGE).

More recently Fodor (1987) has employed LOT to explain other features of cognitive systems, particularly their productivity and systematicity. Productivity refers to the idea that cognitive systems are not bounded: one can always generate new thoughts. Systematicity refers to the idea that having a given cognitive capacity guarantees having certain related capacities. To adapt one of Fodor's remarks, you don't find people who can think Morgan loves Shannon but cannot think Shannon loves Morgan ("mLs but not sLm" in an abbreviated notation). Productivity and systematicity are no accidents, according to Fodor. They fall out of the syntactic structure of the language of thought. Productivity follows from the fact that the rules for building up syntactic structures are recursive, so that one can repeatedly combine composed structures into still larger structures. Systematicity is accounted for by the LOT hypothesis that the brain state that encodes one mental representation (say, "mLs") is a syntactic rearrangement of the brain state that encodes another representation (say, "sLm"). The patterns of possible thoughts and preferences depend on the structure of the language of thought. It is the nature of that structure that regulates the manner in which one thought may lead to or include another.

LOT fits together well with the multiple realizability arguments advanced for functionalism. According to LOT, cognition has nothing directly to do with its species-specific neurobiological embodiment or implementation, but rather concerns processes operating on the common language of thought shared by all entities capable of being in the same cognitive state. Cognition *per se* is not neural; cognition is computation in mentalese. Being in the same cognitive state (say, preferring water) is characterized in LOT as employing the same set of internal symbols and syntactic rules.

Over time, it has been recognized that neither LOT nor functionalism is free of liability. First, we have no inkling as to how or where the syntactically structured mental representations of LOT are stored. In the minds of some critics, that was one huge drawback of Fodor's hypothesis. Second, LOT and functionalism both fail to connect psychology to neuroscience in any experimentally tractable way. For Fodor and Putnam this dissociation between the study of cognition and the study of the brain was not a shortcoming but a strength. Moreover, it fitted well with the *zeitgeist* of cognitive science during the 1970s, which preferred independence from neuroscience. On a practical note, there seemed to be little work in neuroscience that could address cognitive questions such as how people solve problems or learn languages. Through the multiple realizability argument and LOT, Putnam and Fodor seemed to give principled legitimacy to the then current approaches in cognitive science. But by the late 1970s, some philosophers were objecting to the divorce of cognitive science from neuroscience, Paul M. and Patricia S. Churchland foremost amongst them. They tended to continue to endorse a version of the identity theory and to reject LOT (P. S. Churchland, 1986) (see Article 48, LEVELS OF EXPLANATION AND COGNITIVE ARCHITECTURES).

Not all meta-theorizing about cognitive science was done by professional philosophers. We will examine two extremely influential analyses of the project of cognitive science, both of which cohered in many respects with the views of Putnam and Fodor. The first was advanced by neurophysiologist-turned-AI-researcher David Marr in his influential 1982 book *Vision*. In the first chapter he relates his growing disillusionment with the prospects of figuring out how the brain performs cognitive tasks by starting with the response patterns of individual neurons, such as the edge detectors identified by Hubel and Wiesel. He himself had contributed to this project in an important study of the cerebellar cortex (Marr, 1969). Marr's disillusionment was due principally to his coming to realize that discovering the response patterns and wiring diagrams of individual brain parts could only *describe* what was happening in the brain, not *explain* how it accomplished its tasks. He concluded:

> There must exist an additional level of understanding at which the character of the information-processing tasks carried out during perception are [sic] analyzed and understood in a way that is independent of the particular mechanisms and structures that implement them in our heads. This was what was missing – the analysis of the problem as an information-processing task. (1982, p. 19)

Marr went on to argue that in understanding any computational system, such as the brain, two additional levels were needed above the level at which the details of the physical device (the brain) were analyzed. He called his highest level *computational theory* (a label that many have found misleading; it is somewhat akin to Chomsky's notion of competence and might best be called task analysis). Computational theory characterizes the task to be performed, thus answering the questions: What is the goal of this computation? Why is it appropriate? In building an adding machine, for example, the abstractly characterized function of addition is the task we want the machine to carry out. Marr emphasized that the answers to these questions would constrain work at lower levels. The next level, that of *representation and algorithm*, specifies (a) a system of representations (e.g., arabic and roman numerals provide two different representational systems for numbers) and (b) the operations (algorithms) to be performed on them so as to satisfy the function specified in the computational theory (e.g., arithmetic). Marr, like Putnam, emphasized the multiple realizability of computational goals using different representational systems and algorithms. Only at the lowest level, the level of *hardware implementation*, did one turn to the actual physical device and show how it implemented the representations and algorithms from level 2.

It is important to recognize that Marr's levels identify three different ways to analyze the same system; they do not characterize levels of organization in nature, which typically are related in a part–whole fashion. Moreover, while he emphasized the differences between levels and noted the multiple realizability of systems satisfying higher levels in the analysis, he also recognized a loose coupling of the levels: the analysis advanced at any given level constrained analyses at the others. Finally, while Marr is sometimes portrayed as proposing that we work solely from the highest level down, he clearly emphasized constraints coming from the bottom up as well and employed them in his attempt to explain visual processing. His concern was ultimately to figure out what an adequate explanation would look like, which brought him to the claim that it would have to provide analyses at each of these three levels.

While Marr's analysis of levels is in some respects compatible with Putnamian functionalism (but recognizes more constraints between levels than Putnam indicated), the other case of meta-theorizing by nonphilosophers, Newell and Simon's *Physical Symbol System Hypothesis*, shared Fodor's emphasis on symbolic processing as a key to understanding cognition. The hypothesis (as presented in Newell, 1980) states that a physical symbol system satisfies the necessary and sufficient conditions for exhibiting "general intelligent action" (p. 170). The kind of system envisaged by Newell and Simon is a universal system that can carry out any mapping from input states to output states and does this by operating on symbols. What makes something a symbol is its designation or reference:

> The most fundamental concept for a symbol system is that which gives symbols their symbolic character, i.e., which lets them stand for some entity. We call this concept designation, though we might have used any of several other terms, e.g., reference, denotation, naming, standing for, aboutness, or even symbolization or meaning. . . .
>
> Let us have a definition:
>
> > *Designation*: An entity X designates an entity Y relative to a process P, if, when P takes X as input, its behavior depends on Y.
>
> There are two keys to this definition: First, the concept is grounded in the behavior of a process. Thus, the implications of designation will depend on the nature of this process. Second, there is action at a distance . . . This is the symbolic aspect, that having X (the symbol) is tantamount to having Y (the thing designated) for the purposes of process P. (p. 156)

Consider, by way of analogy, a map that represents Boston. If the map has the right resources for designation, and someone who knows how to read maps applies that process appropriately, then Boston-on-the-map can stand for – is tantamount to – Boston in Massachusetts.

What philosophers give, they also can take away. Dreyfus and his fellow Berkeley philosopher John Searle were especially critical of approaches to cognition like those of Putnam and Fodor as well as Newell and Simon. Both these philosophers denied that cognition consists in syntactic processing of symbols alone. Dreyfus (1972/1979) argued that much natural cognition is not representable in a symbolic code, but is, in some sense, out in the environment. To do cognitive science, according to Dreyfus, one must work with embodied systems, not syntactic processors – a theme some cognitive scientists embraced in the 1990s (see Article 39, EMBODIED, SITUATED, AND DISTRIBUTED COGNITION).

Searle was a skeptic of his own kind. Symbolic codes, Searle claimed, are not even cognitive; they lack genuine *cognitivity* (what Searle calls *intrinsic intentionality*, the property of being about something). Real mental states possess INTENTIONALITY in themselves; your belief that there is a cold beer in the refrigerator has that content in itself and does not depend on an external interpreter to assign it that content. Codes possess only derived intentionality (based on outside interpretation). Searle (1980) argued for these claims by constructing a thought-experiment which he dubbed the "Chinese Room." In the thought-experiment he plays the role of a computer that is programmed to answer in Chinese questions asked in Chinese, much in the manner of trying to pass a Turing test. Searle imagines himself in possession of a book of syntactic rules which tells him how to replace strings of Chinese characters (in questions) with other strings of Chinese characters (in answers). Outside observers may believe that the computer/person in the room understands Chinese; the *room* will pass the Turing test. But the rule book is a grammatical manual, not a bilingual dictionary. It does not tell the denizen of the room what the characters mean. So Searle contends that he could execute all the syntactic manipulation proposed by a theory like LOT, yet not understand – not cognize in – Chinese. Therefore computation over a system of meaningless symbols is insufficient to account for cognition. Such symbolic codes are not cognitive.

The objections of Dreyfus and Searle, to the extent that they are valid, may threaten not cognitive science in general, but rather those cognitive science research programs that are wedded to an *exclusively* computational view of cognition. One may concede that contemporary computers are purely syntactic engines that lack intentionality. Hence, they are not cognitive. But what this means is that to answer Searle, one must develop some way to *ground* the symbols of a system so as to give them genuine content. Some theorists have suggested that it is the way in which the symbols are acquired by the system (e.g., through a connectionist learning process in which the structure of the representation is acquired through interaction with the objects to which they refer) that makes them intentional (Harnad, 1990). Another strategy is to deny that our mental states enjoy intrinsic intentionality – they only seem to do so because we have the ability to use yet other symbols, such as words in our natural language, to specify their referents. To accommodate Dreyfus, cognitive science may have to take seriously that our cognitive systems are embodied and situated in the world, a theme that, as we have noted, a variety of contemporary researchers are pursuing.

2.7 *Getting an identity*

In the modern world, intellectual developments such as those we have been discussing require institutions which frame and support their activities. These institutions include departments or other administrative units in universities, as well as journals and professional societies. We turn now to the process by which cognitive science developed its institutional identity towards the end of its period of initial maturation.

2.7.1 *Cognitive science centers: the legacy of the Sloan Foundation*

Academic researchers are generally hired into departments, and the structure of departments plays a critical role in their ongoing life. Departments allocate the important commodities, such as space, money, and teaching assignments. They control the curriculum for both undergraduate and graduate education, and thus determine the intellectual frameworks and tools that new investigators will be prepared to employ. Even more important, they make important personnel decisions about whom to hire and whom to tenure.

As cognitive scientists from different disciplines began to explore their common interests, they sought an institutional structure in their home universities. Sometimes new academic pursuits result in new departments, and this has happened at a few locations for cognitive science (e.g., at UCSD and Johns Hopkins University). But creating new departments is not always the best path of development for new areas of inquiry to take, especially when they bridge existing disciplines rather than offer new methodologies for exploring new domains. First, departments cost money, and universities are often reluctant to use scarce resources in this way. Second, moving into a new department can isolate a researcher from others who share a common disciplinary foundation (e.g., it could isolate a cognitive psychologist from other psychologists, many of whom employ similar research tools and theoretical frameworks).

A solution to the desire of cognitive scientists both to retain affiliation with their home disciplines and to affiliate with other cognitive scientists from different departments is to create a center. The Harvard Center for Cognitive Studies, discussed above, was a prototype. The creation of such centers is not unique to cognitive science; universities have created centers (sometimes called *committees* or *programs*) in a variety of interdisciplinary areas, including women's studies, African-American studies, history and philosophy of science, evolutionary biology, and materials science. In some instances, these programs or centers become virtual departments, running their own degree programs, tenuring their own faculty, and so forth. But generally they serve more as meeting grounds, by sponsoring colloquia and postdoctoral fellowships; sometimes they also develop a curriculum comprised of courses from various departments as well as new, explicitly interdisciplinary courses. Instead of offering their own degrees, at least at the Ph.D. level, they may offer certificates which complement a graduate degree in one of the existing departments.

The center model was quickly adopted in cognitive science. As of 1983, the following North American universities had all developed centers for cognitive science: Brown University, Carnegie–Mellon University, the University of Chicago, the University of Colorado, Cornell Medical School, the University of Illinois at Urbana, the University of Massachusetts at Amherst, Massachusetts Institute of Technology, the University of Michigan, the University of Pennsylvania, the University of Pittsburgh, Princeton University, the University of Rochester, Rutgers University, Stanford University, the University of Texas at Austin, the University of Western Ontario, Yale University, and four campuses of the University of California (Berkeley, Irvine, Santa Barbara, and San Diego). The initiative for creating these centers in part arose from faculty members who realized that many of their closest intellectual alliances involved faculty from other departments. But another powerful incentive behind the creation of many of these centers was financial. Starting in 1977, the Alfred P. Sloan Foundation provided substantial grants to universities to help them establish institutional structures supporting research in cognitive science. In the Foundation's annual report for 1977, this initiative was characterized as follows:

> Scientists in several disciplines, including psychology, linguistics, computer science, and neuroscience, believe that by pooling their diverse knowledge they can achieve a more advanced understanding of human mental processes such as memory, perception, and language. The Particular Program in Cognitive Sciences seeks to accelerate this cross-fertilization at some of the most promising centers for such research. (p. 48)

During its first two years of funding cognitive science, the Sloan Foundation focused on activities that were *exploratory*: "intended to enable institutions with varying disciplinary strengths to consider the kinds of coordinated research and training programs which they might develop in the future" (report for 1978, pp. 55–6). In the initial year of support, grants primarily supported workshops and visiting scientists. The largest of these, to UCSD, supported a number of postdoctoral fellowships and a major conference in summer 1979 which was to become the first conference of the Cognitive Science Society (see below).

In the course of identifying its mission in cognitive science, the Sloan Foundation commissioned a report on the state of the art in cognitive science from a committee of advisers from the cognitive sciences. In its report of October 1, 1978 (excerpted in Pylyshyn, 1983, appendix, p. 75), the committee argued that an "autonomous science of cognition has arisen in the past decade" on the basis of "the richly articulated pattern of interconnection among [the] subdomains" of neuroscience, computer science, psychology, philosophy, linguistics, and anthropology. They presented the pattern of connections in a diagram, in which six nodes represent the contributing disciplines, and connections between them represent interdisciplinary endeavors (figure I.10). The labeled connections represent well-defined areas of cross-disciplinary inquiry. For example, psycholinguistics joins the research methods and conceptual underpinnings of experimental psychology to those of linguistics. The four dotted lines represent interdisciplinary connections that had not yet taken form. The committee also noted that one can identify some three-way connections, such as the cooperation among psychology, linguistics, and philosophy in the study of language. The authors then offered a statement of the common research objective that links all the disciplines of cognitive science: "to discover the representational and computational capacities of the mind and their structural and functional representation in the brain" (p. 76).

With the vision of what cognitive science should be, and on the basis of the initial exploratory grants, the Sloan Foundation moved into a second phase, "the establishment of formal training programs in the emerging discipline of cognitive science" (report for 1979, p. 49). In that year they made six grants in the $400,000 to $500,000 range to UCSD, MIT, Carnegie–Mellon University, the University of Pennsylvania, the University of Texas, and Yale University. A brief sketch of the initiative at two of these institutions gives a sense of the scope of the cognitive science centers that Sloan was supporting. Sloan characterized the interests of the members of the Center for Human

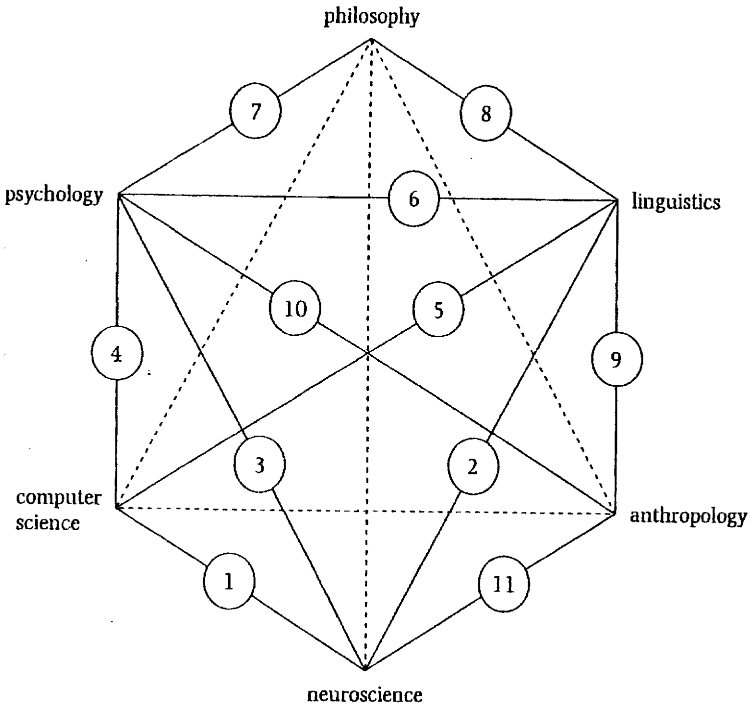

Figure I.10 Representation of the domains and subdomains of cognitive science according to the 1978 Sloan Report on Cognitive Science. Each of the six contributing disciplines is represented by a node. The labeled connections between nodes represent formally organized interdisciplinary collaborations: (1) cybernetics, (2) neurolinguistics, (3) neuropsychology, (4) simulation of cognitive processes, (5) computational linguistics, (6) psycholinguistics, (7) philosophy of psychology, (8) philosophy of language, (9) anthropological linguistics, (10) cognitive anthropology, (11) evolution of the brain. The dotted lines represent collaborations which had not yet formally developed.

Information Processing at UCSD as including "formal analyses of linguistic systems, neural mechanisms of cognitive functioning, and anthropological investigation of human belief systems." With the new grant, "UCSD is now instituting a postdoctoral training program drawing upon insights gained from computational methods, theoretical psychology, neuroscience, and related experimental techniques" (pp. 49–50). MIT, by comparison, is characterized as having "rich resources in linguistics, philosophy, artificial intelligence, speech research, cognitive psychology, and neuroscience. Its administration has supported the creation of a Center for Cognitive Science which will draw upon all of those resources and provide some new ones" (p. 50). The following year Sloan made seven additional grants of similar size and focus to the University of Chicago and University of Michigan (in a inter-university collaboration), Stanford University, Brown University, the University of Massachusetts at Amherst, Cornell Medical College (in conjunction with Rockefeller University), and UC–Irvine. The grant to Cornell Medical College is noteworthy, since it was the only grant to place primary emphasis on neurophysiological information (by seeking to use analyses of brain damage to make inferences about cognitive functions in non-brain-damaged individuals).

The following year, 1981, Sloan took its most ambitious step by committing amounts ranging from $500,000 to $2.5 million to the "establishment at each participating institution of an identifiable, self-sustaining center, institute, department, or other administrative entity where a continuing program of research and training will be conducted in cognitive science" (report for 1981, p. 12). The following list identifies the institutions selected and the amounts they received:

Massachusetts Institute of Technology	$2.5 million
University of California, Berkeley	$2.5 million
Carnegie–Mellon University	$1.0 million
Stanford University	$1.0 million
University of Pennsylvania	$1.0 million
Cognitive Neuroscience Institute	$0.5 million
University of California, Irvine	$0.5 million
University of Rochester	$0.5 million
University of Texas, Austin	$0.5 million

While most of these schools and institutes were ones previously supported by Sloan (the Cognitive Neuroscience Institute was a continuation of the program at Cornell Medical College under Gazzaniga), there were some interesting changes. UCSD, which had been heavily funded in the two previous phases, was not included (perhaps because it already had a relatively stable institutional base in place), whereas the University of Rochester, in a program under the direction of philosopher Patrick Hayes, was newly selected (perhaps in part as a result of the University's express commitment to cognitive science in the form of an established chair).

Another relative newcomer to Sloan support was the University of California at Berkeley, which received an exploratory grant only after other institutions were already being selected for training grants. The Sloan Foundation explained its grant to Berkeley in the following way (it helps to keep in mind that Searle and Dreyfus taught at Berkeley):

> The unorthodox approach researchers at Berkeley take to cognitive science – sometimes called the "Berkeley approach" – is characterized by a vigorous skepticism toward some of the basic assumptions underlying research at other cognitive science centers. Berkeley philosophers, for example, have been extremely critical of some prominent claims made for artificial intelligence. There is no reason, they assert, to believe that machine intelligence can ever approximate human intelligence or achieve anything like human consciousness. Similarly, the Berkeley linguistics department has become a center of opposition to transformational theory as developed at M.I.T. A major attempt has been made at Berkeley to develop a theory of language that is part of a general theory of action, an approach that differs sharply from the more analytic attempts of transformational linguists to segregate linguistic competence from the other psychological mechanisms involved in language use. Because the field of cognitive science is so new and research is still so rudimentary, it seems to us and to our outside advisory committee only prudent for the Foundation to place some bets on those who dissent from the majority views. (Report for 1983, p. 16)

After ten years, in which it provided approximately $17.4 million for direct costs, the Sloan Foundation brought its initiative in cognitive science to a close. During part of this period (1982–4) a second foundation, the System Development Foundation, was also pursuing an initiative in areas related to cognitive science. Focusing on computational linguistics and speech, it invested a total of $26 million. The largest single grant went to support the Center for the Study of Language and Information (CSLI) at Stanford University. CSLI was created in 1983, as a research institute connecting faculty at Stanford with researchers at SRI and Xerox-Parc in collaborative inquiry into human–computer interaction, language processing, and reasoning. The participants were drawn primarily from linguistics, philosophy, and computer science. One of the best-known projects of CSLI was initiated by philosophers Jon Barwise and John Perry: they developed a model of situation semantics which analyzed the meaning of mental states and linguistic utterances in terms of their relations to situations in the actual world.

The funding by both the Sloan Foundation and the System Development Foundation helped to build the institutional base for cognitive science, but since both foundations committed support for only a relatively short period, a serious question arose as to sources of future funding. In anticipation of this problem, on March 30 and April 1, 1985, a workshop was held under the sponsorship of the National Science Foundation and the System Development Foundation. Reflecting the orientation of the System Development Foundation, this workshop did not adopt the term *cognitive science* but rather referred to the *study of information, computation, and cognition*. It nonetheless identified the same contributing disciplines as had the Sloan Foundation, with the notable omission of neuroscience. The workshop assessed the current state of funding in the cognitive sciences. While some cognitive science research was supported by various agencies within the Department of Defense, especially the Office of Naval Research, the most dependable source of funding was the National Science Foundation. The report advocated a significant increase in funding through NSF and other federal agencies.

2.7.2 *The journal and the society*

While it might have seemed easy for researchers sharing an interest in cognition to communicate with one another, this actually became one of the most serious challenges for cognitive science. Researchers trained in different fields conceived of their inquiries differently, used different research tools, spoke in different vocabularies, and read different literatures.

A study that Zenon Pylyshyn conducted in the late 1970s, the period when cognitive science was already taking form, revealed a serious communication gap (Pylyshyn, 1983). He surveyed citation patterns within artificial intelligence and cognitive psychology publications, sampling a total of 528 references from the proceedings of the International Joint Conference on Artificial Intelligence (IJCAI) in 1977 and from two years of the journal *Artificial Intelligence*. Of the references that could be categorized according to the discipline of origin, 300 of these were to other papers and books in artificial intelligence; only 35 were to books or journals in psychology and 7 to journals in linguistics. An important feature of the citations is that many were to unpublished technical reports, which often would not be easily available to those outside the field. The situation in psychology was similar. Of the 1,200 citations that Pylyshyn sampled in *Cognitive Psychology*, *Cognition*, and *Memory and Cognition*, nearly 1,000 were to books or journals in psychology, with a majority of the journal references to journals in general experimental psychology that did not emphasize a cognitive or information processing perspective. Only 50 citations were to AI sources, 70 to linguistics articles, and 16 to papers in philosophy. The majority of the references in the latter two categories appeared in *Cognition*, which explicitly solicited interdisciplinary work.

An obvious reason for the relative paucity of references across cognitive science disciplines was that, except in specially arranged meetings, researchers in one area would not easily encounter the work of those in related areas. Generally a cognitive psychologist or linguist would not subscribe to an AI journal, and vice versa. They would also not attend each others' professional meetings, unless explicitly invited to give a talk. For one thing, there was a significant cost to subscribing to journals or registering for meetings outside one's field. In addition, each profession had its own network of people who gave talks and wrote papers intended for each other. It was not easy for an outsider to gain admittance. If there were to be real interdisciplinary communication, cognitive science clearly needed to establish forums devoted to this goal.

Accordingly, in 1977 a new journal called *Cognitive Science* was founded under the editorship of Roger Schank, Allan Collins, and Eugene Charniak. With their diverse backgrounds, the three editors seemed to represent the interdisciplinary mix that was beginning to characterize cognitive science. As noted above, Schank's Ph.D. was in linguistics, but his subsequent research led him into computer science and AI. Allan Collins received his Ph.D. in psychology from Michigan in 1970 and then held a research position at Bolt Beranek and Newman (a Boston-area high technology firm which employed a number of cognitive science researchers). Following his collaboration with Quillian (discussed above), at BBN Collins developed research approaches that belonged to cognitive science itself rather than any traditional discipline, and then joined Schank in moving towards educational applications (especially through the Institute for the Learning Sciences). Eugene Charniak's Ph.D. was in Computer Science from MIT, and at the time he was at the Institut pour les études semantiques et cognitives in Geneva, Switzerland. But the mix was not quite as eclectic as it might seem. While Schank was a linguist by training, his interest was in systems for natural language understanding and informal reasoning. Moreover, he was already an outspoken critic of Chomsky; accordingly, the MIT orientation to linguistics, which had played a significant role in the genesis of cognitive science, was not represented. Indeed, the banner underneath the title of the journal read "A multidisciplinary journal of artificial intelligence, psychology, and language," referring to language rather than linguistics.

The first issue began with an editorial by Allan Collins entitled "Why Cognitive Science" in which he described the converging interest in cognition involving a number of disciplines and characterized cognitive science as a new discipline with a distinctive view of natural and artificial intelligence: "This view has recently begun to produce a spate of books and conferences, which are the first trappings of an emerging discipline. This discipline might have been called applied epistemology or intelligence theory, but someone on high declared it should be cognitive science and so it shall. In starting the journal we are just adding another trapping in the formation of a new discipline" (p. 1).

A fuller indication of the orientation of the new journal is provided in box I.2, which shows the table of contents of the first volume, together with the names of the authors and their institutional affiliations. The authorship of papers reveals a significant mix

Box I.2 Table of Contents from first volume of *Cognitive Science*

Volume 1, Number 1

Daniel G. Bobrow (Xerox Palo Alto Research Center) and Terry Winograd (Computer Science, Stanford University). "An Overview of KRL, a Knowledge Representation Language"

Wendy Lehnert (Computer Science, Yale University). "Human and Computational Question Answering"

Andrew Ortony and Richard C. Anderson (Psychology, University of Illinois at Urbana–Champaign). "Definite Descriptions and Semantic Memory"

Ira Goldstein and Seymour Papert (MIT AI Labs). "Artificial Intelligence, Language, and the Study of Knowledge"

Volume 1, Number 2

John R. Anderson (Psychology, Yale). "Induction of Augmented Transition Networks"

Jerome A. Feldman (Computer Science, University of Rochester) and Robert F. Sproull (Xerox Palo Alto Research Center). "Decision Theory and Artificial Intelligence II: The Hungry Monkey"

B. Bhaskar and Herbert A. Simon (Psychology, Carnegie–Mellon University). "Problem Solving in Semantically Rich Domains: An Example from Engineering Thermodynamics"

M. J. Steedman (Psychology, University of Sussex). "Verbs, Time, and Modality"

Volume 1, Number 3

Yorick Wilks (Artificial Intelligence, Edinburgh). "What Sort of Taxonomy of Causation Do We Need for Language Understanding?"

Stephen M. Kosslyn and Steven P. Shwartz (Psychology, The Johns Hopkins University). "A Simulation of Visual Imagery"

J. R. Hayes, D. A. Waterman, and C. S. Robinson (Psychology, Carnegie–Mellon University). "Identifying the Relevant Aspects of a Problem Text"

Chuck Rieger (Computer Science, University of Maryland). "Spontaneous Computation in Cognitive Models"

Volume 1, Number 4

Eugene Charniak (Institut pour les études semantiques et cognitives, University of Geneva). "A Framed Painting: The Representation of a Common Sense Knowledge Fragment"

James A. Levin and James A. Moore (USC Information Sciences Institute). "Dialogue-Games: Metacommunication Structures for Natural Language Interaction"

Roger C. Schank (Computer Science, Yale University). "Rules and Topics in Conversation"

of psychologists and AI researchers. In terms of citation statistics, Pylyshyn (1983) reported that *Cognitive Science* represented a far greater amount of cross-disciplinary integration than the other journals he examined: "Out of 331 citations of the last two years, 110 were judged to be clearly psychological (31 books and 79 journal articles in psychology, with 29 of these in cognitive psychology); 55 were artificial intelligence papers; 14 were articles in computer science journals; 50 were citations of other articles in *Cognitive Science* itself; 40 were journal articles in philosophy and logic; 26 were linguistics papers; 7 were neuropsychological papers; and the remaining 36 citations were distributed among a variety of areas, including library science, education, business, anthropology, and book reviews" (p. 72). What these statistics do not reveal is whether individual authors were citing papers outside their own disciplines, revealing integration at the level of individual research, or whether *Cognitive Science* was simply providing a common forum.

The original conception of *Cognitive Science* emphasized psychology and AI. Such disciplines as neuroscience and philosophy were not expressly included in its scope, and linguistics of the Chomskian sort was not represented. But *Cognitive Science* was not the only new journal trying to reach an interdisciplinary community in cognitive science. In 1977, the same year that *Cognitive Science* was first published, the Society for the Interdisciplinary Study of the Mind began to publish its own newsletter, the *SISTM Quarterly*. In 1979–80 *SISTM Quarterly* incorporated the *Brain Theory Newsletter*, and took the name *Cognition and Brain Theory: The Newsletter of Philosophy, Psychology, Linguistics, Artificial Intelligence, and Neuroscience*. As the term *newsletter* suggests, it was initially published as a stapled, 8.5 by 11 inch pamphlet and was distributed directly by Vassar College where Martin Ringle was the editor. In addition to news notices, it published abstracts, bibliographies, and articles. Publication of *Cognition and Brain Theory* was transferred in 1981 to Lawrence Erlbaum Associates, where it took on the format of a standard journal. Both as a newsletter and as a journal, *Cognition and Brain Theory* tended to give greater coverage to philosophy and neuroscience than did *Cognitive Science*. In its first two issues in journal format, *Cognition and Brain Theory* included papers by linguist Noam Chomsky, neuroscientists Michael Arbib and Karl Pribram, neuropsychologists Daniel Bub and Harry Whitaker, psychologist Dedre Gentner, AI researchers Nils Nilsson and Roger Schank, and philosopher William Lycan. After 1984, *Cognition and Brain Theory* merged with *Cognitive Science*, but with a commitment from *Cognitive Science* to broaden its scope and publish more neuroscience and philosophy.

In 1978, one year after *Cognitive Science* and *Cognition and Brain Theory* were established, Stevan Harnad, then a graduate student at Princeton, created yet another new journal, *Behavioral and Brain Sciences* (*BBS*), published by Cambridge University Press. *BBS* adopted a distinctive format of primary papers (called *target articles*), followed by a dozen or more short commentaries. It invited participation from the entire cognitive science community in writing both target articles and commentaries. Target articles were selected with an eye to their potential for stimulating multidisciplinary dialogue. For example, in 1980 *BBS* published John Searle's "Minds, brains, and programs," which advanced his Chinese room critique of computational accounts of mind (see above), together with responses by psychologists such as Robert Abelson and Bruce Bridgeman, philosophers such as Ned Block, Daniel Dennett, Jerry Fodor, and William Lycan, neuroscientists such as John Eccles and Benjamin Libet, and AI researchers such as Zenon Pylyshyn, John McCarthy, Marvin Minsky, and Roger Schank.

At this point (the late 1970s), there were at least three journals supporting the emerging interdisciplinary field of cognitive science. Two new academic presses soon created similar opportunities for publishing book-length works. First, Lawrence Erlbaum established Lawrence Erlbaum and Associates, which quickly amassed a strong list of books in cognitive psychology and gradually expanded into other cognitive science fields such as artificial intelligence and linguistics. Second, Harry and Betty Stanton established a small press, Bradford Books, which was subsequently acquired by MIT Press but retained its separate editorial process and imprimatur. Bradford's initial focus was on philosophy texts such as *Brainstorms* by Daniel Dennett and *Knowledge and the Flow of Information* by Fred Dretske, and it gradually expanded into the cognitive sciences more generally.

But there was not yet any professional organization. One of the most important functions performed by professional organizations is to sponsor annual meetings at which researchers can report on work they are doing and learn about new theoretical ideas and empirical results obtained by others. Discussions, both formal ones after presentations and informal ones in hallways and lounges, serve as a powerful stimulus to new thoughts and help to ratify promising ideas. Conferences were even more important before the era of rapid electronic communication than they are today, since circulation of early reports of research was generally limited to small groups of fellow researchers. As we have noted, conferences were arranged in cognitive science without the sponsorship of a professional organization, but they depended upon the interests and resources of the particular researchers who organized them. Once a professional organization is in place, it raises funds, both through dues and grants, to cover the costs of planning a meeting, and deputizes people to carry out the necessary work (arrange meeting locations, plan a program, and so on).

It was in the course of planning what was to be just one more individually arranged conference that the idea of the Cognitive Science Society took form. The newly developed interdisciplinary program in cognitive science at UCSD, in the context of its grant from the Sloan Foundation (see above), set the La Jolla Conference on Cognitive Science for August 13–16, 1979. The goal of the conference was to assess the state of the art in cognitive science by having ten leading figures in the profession each address "some of the hopes, aspirations, and critical issues that face the development of a cognitive science" (Norman, 1981, p. v). As Norman relates, the goal of the conference was nothing less than to define *cognitive science*:

> It was to be the "defining meeting," the meeting where many of those concerned with the birth of Cognitive Science could record its origins, speak of its hopes, and chart its course. We knew these aspirations to be unrealistic, but did not let that knowledge deter us. The speakers at the conference – the contributors to this volume – all work within the sibling disciplines that comprise Cognitive Science. All were charged with the task of presenting broad, overview statements of their views, statements that would last beyond the year of the conference and that would help set the definition of the field, statements that would prove useful in the initial stages of the discipline and that would provide examples of what we are, what we wish to become, and even what we should not be. (Ibid.)

Five of the eleven speakers came from AI (Herbert Simon, Allen Newell, Marvin Minsky, Roger Schank, and Terry Winograd), two from psychology (Donald Norman and Philip Johnson-Laird), one from neuroscience (Norman Geschwind), one from linguistics (George Lakoff), and two from philosophy (John Searle and Mark Johnson, coauthor of the paper with Lakoff). Norman notes that there was concern about the omission of "what might be called the 'MIT school of linguistics'," but that none of the three investigators who were invited were able to attend. One feature of defining activities is that they are selective: they emphasize certain perspectives at the expense of others. This was true of the La Jolla meeting. AI was clearly the dominant discipline, and it has continued to play a fundamental role in the Cognitive Science Society, whereas Chomskian approaches received less attention. Nonetheless, the Cognitive Science Society grew rapidly. Its meetings now consist of a large number of paper sessions scheduled in parallel, as well as large poster sessions, and attract about a thousand attendees. In addition to publishing a large volume of papers from its annual meeting, in 1980 it became the sponsoring organization for the journal *Cognitive Science*.

3 Identity crises: 1985–1999

The focus of cognitive science had narrowed in the period 1960–1985 from its earlier breadth. Beginning in the 1980s, it regained that breadth and more by expanding in two directions: vertically into the brain and horizontally into the environment. These expansions, however, have induced an identity crisis for the developing cognitive science enterprise. During its phase of initial maturation the scope of cognitive science was reasonably well demarcated, with an accepted framework of representations and computations over them defining the accepted explanatory approach. The movement outwards into the environment and downwards into the brain has given rise to other models for cognitive science. Many of these developments are covered in detail elsewhere in the Companion; accordingly, our interest here is simply to place them in historical context.

3.1 Rediscovering neural networks

After a flourish of research in the initial, revolutionary period of cognitive science (see section 1.3.5), research into artificial neural networks largely disappeared from artificial intelligence and cognitive science after 1970. One reason for this was that, like traditional AI researchers, neural network researchers such as Rosenblatt often made claims that they could not substantiate. A second reason was that, while neural networks' forte is pattern recognition, an activity that seems most suited to analyzing sensory input, much of the interest in cognition was focused on higher cognitive processes such as reasoning and problem solving. A third reason for the decline of neural network research was a serious limitation found in neural networks of the sort Rosenblatt was developing. As discussed in section 1.3.5, his greatest achievement was to develop and prove the effectiveness of a learning procedure for modifying connection weights in networks with one layer of connections. But there were a large number of problems which could not be solved by such a simple perceptron.

Rosenblatt was aware of this limitation, and was in the process of exploring networks with multiple layers of connections when Minsky and Papert published an extremely influential book entitled *Perceptrons* in 1969. The single-layer perceptron was made the object of detailed mathematical study. One of the results the authors presented was a proof that such networks could not perform certain computations, such as *exclusive or* (*XOR*). (A XOR B is true if and only if just one of the two propositions, A or B, is true. It contrasts with the usual logical connective *or*, which is true when at least one of the two propositions is true.) Although Minsky had himself initially worked on neural network models, he became disaffected from that approach and adopted the symbolic framework instead. *Perceptrons* was broadly perceived as a fundamental attack on the whole neural network approach. It is less clear, however, whether it was responsible for driving away funding for neural networks generally, or was rather only one sign that the approach had fallen into disfavor.

Whatever the causes, neural network research did not show the same growth pattern during the 1970s as did the symbolic approach. Nonetheless, some important research continued to be carried out by researchers such as James Anderson, Teuvo Kohonen, Christoph von der Malsburg, and Stephen Grossberg. At the beginning of the 1980s interest began to revive. In part this was due to emerging dissatisfaction with some of the perceived limits of symbolic models, such as their failure to degrade gracefully as the systems were partially damaged, restricted ability to generalize to new cases, and inefficiency as models grew more complex. An important testimony to renewed interest in the new approach (now often referred to by the terms *connectionism* and *parallel distributed processing* as well as *neural networks*) was the publication of *Parallel Models of Associative Memory* in 1981, edited by Geoffrey Hinton and James Anderson. It contained papers from a conference held at UCSD in June, 1979, which included neuroscientists, cognitive psychologists, AI researchers, mathematicians, and electrical engineers, all exploring the neural network approach. Many of the participants were recent Ph.D. recipients whose involvement breathed new life into the neural network program.

Two important papers in the 1980s further fueled the rejuvenation of interest in neural networks. The first, by John Hopfield (1982), provided a very clear exposition of how computation was possible in neural networks; it was influential, though, largely because of Hopfield's status as a leading physicist:

> John Hopfield is a distinguished physicist. When he talks, people listen. Theory in his hands becomes respectable. Neural networks became instantly legitimate, whereas before, most developments in networks had been the province of somewhat suspect psychologists and neurobiologists, or by those removed from the hot centers of scientific activity. (Anderson and Rosenfeld, 1988, p. 457)

The second paper, by David Rumelhart, Geoffrey Hinton, and Ronald Williams (1986), cracked the problem of how to train multilayered networks. The extra layer(s) of connections together with the new learning procedure enabled these networks to solve the XOR problem and others that had figured in Minsky and Papert's critique of Rosenblatt. The learning procedure, known as *back-propagation*, took Rosenblatt's procedure for one-layer networks as a starting point. That is, the network would use its current weights to respond to an input pattern, and the difference between the actual output and the target output for each unit would constitute an error signal which would be passed back through the network. Weights were adjusted in a direction that would reduce the error on that input pattern. Rumelhart et al. found a way to repeat this process for additional layers of connections. (See Article 16, MACHINE LEARNING.) Subsequent research established that networks with two layers of connections and sufficient numbers of hidden units could be trained in this way to compute any computable function, thereby rendering them equivalent in computational power to universal Turing machines or digital computers. In 1986 Rumelhart, James McClelland, and the PDP Research Group (based at UCSD) published a two-volume collection entitled *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, which quickly became the bible of the connectionist movement.

Neural networks provided an answer to several of the sources of dissatisfaction that were emerging with symbolic models. When networks are damaged, they do not simply crash; rather, their performance generally degrades gradually (e.g., responses will become increasingly less accurate). The same weights that enable the networks to master training cases enable them to generalize to new cases. In addition, research on neural networks introduced different ways of thinking about mental representations. Representations in networks are often distributed; that is, the representation of a single item will involve more than one unit, and a given unit will participate in the representation of more than one item. Individual units then encode information at a *sub-symbolic* level (Smolensky, 1988). For example, a set of words might be represented by distinctive but overlapping patterns of activation values across ten units. The values could be arbitrary, or the units could be engineered to correspond to phonemic or semantic features. Distributed representations turn out to be robust not only against damage but to permit interesting generalizations to new cases.

The re-emergence of neural networks in the 1980s resulted in considerable rancor in the cognitive science community. To theorists such as Jerry Fodor and Zenon Pylyshyn, the return of neural networks represented a relapse into associationism, which they contended had already been shown to be inadequate by such proofs as Chomsky's arguments about the insufficiency of statistical models to account for grammatical structure in language. Fodor and Pylyshyn (1988) presented a broad theoretical critique of connectionism, whose centerpiece was the objection that since connectionist networks did not employ a representational system with combinatorial syntax and semantics, they could not account for the productivity and systematicity of thought. We encountered these notions above in discussing Fodor's arguments for a language of thought. Far from surrendering in the face of such objections, however, connectionists have issued a number of vigorous responses (see Article 48, LEVELS OF EXPLANATION AND COGNITIVE ARCHITECTURES, and Article 52, RULES). Neural networks have become increasingly influential with cognitive researchers, many of whom regard them as simply part of the tool-kit for modeling cognition, rather than a call to arms in a theoretical battle.

What made neural network models radical for some cognitive scientists was that they seemed to do away with syntactically structured representations as the currency within a cognitive system. An even more radical approach has emerged more recently under the banner *dynamical systems theory* (*DST*). The DST approach has sought to apply to cognitive phenomena the same type of dynamical equations that have proved successful in physics, beginning with Newton's formulation of the force laws. The key to the approach is to develop mathematical relations, often nonlinear, among parameters characterizing features of a cognitive system (sometimes relating these to parameters characterizing features of the environment). The nonlinearities in the equations often lead to complex patterns of change, but methods for obtaining geometric representations make it easier to track the behavior of such systems (Port and van Gelder, 1995). Among the more radical claims of some DST advocates has been a repudiation of (a) the framework of REPRESENTATION AND COMPUTATION that has been basic to most cognitive science and (b) the general strategy of explaining cognitive processes in terms of different components performing different sub-tasks (see section 4).

Another theoretical perspective that emerged at the end of the 1980s and derived in part from neural network research is research on *Artificial Life* (Langton, 1995). This rubric incorporates a wide range of approaches which generally share the perspective of taking living systems, not just cognitive systems, as their focus. One inspiration for it was the discovery that one could develop computer models that exhibited a broad range of behaviors characteristic of living organisms. Another was the discovery by investigators such as Rodney Brooks that one could develop robots which could negotiate environments without a centralized system for planning movement. A final inspiration was the discovery by John Holland of computer algorithms (*genetic algorithms*) which used simulated evolution to develop new programs that were better adapted to the tasks they faced than were their ancestors. While often themselves critical of cognitive science, Artificial Life researchers have developed a range of new ideas that have percolated through cognitive science. (For more on the approaches introduced in this section, see Article 38, CONNECTIONISM, ARTIFICIAL LIFE, AND DYNAMICAL SYSTEMS).

3.2 Rediscovering the brain: cognitive neuroscience

During 1960–80, most cognitive scientists only occasionally considered neuroscience and the brain as they went about the business of developing and testing theories and models. Reasons for ignoring neuroscience were both pragmatic and principled. On the pragmatic side was the fact that the questions asked and the tools used in much neuroscience research were remote from the inquiries being conducted in cognitive science, especially in cognitive psychology. (An exception is that many psychologists and linguists kept an eye on the intriguing flow of results from neurolinguistics.) On the principled side, the very metaphor that enabled artificial intelligence to play such a major role in cognitive science, the computer model, also provided a framework for minimizing the relevance of neuroscience. The idea was that the relation between psychology and neuroscience was like the relation between hardware and software. Hardware is required for any software to run – to adapt the language of Putnamian functionalism, the hardware implements the functional roles constitutive of the software – but the details of the hardware generally do not matter. One can run the same high-level program on different hardware systems, with no differences in what the software does. (This requires the appropriate lower-level software – a compiler or interpreter – for each hardware system. There may be differences in how long the program takes to run, what memory registers get used, and so on, but these can be dismissed as *implementation* details.) If mental processes constitute something like the software responsible for behavior, while the brain is the hardware, then it would seem that cognitive science could proceed without concern for neuroscience.

These reasons for neglecting the brain began to be challenged in the 1980s. On the one hand, the computer metaphor began to lose its hegemony. In part this was because connectionist (neural network) models of cognitive function re-emerged, thereby undercutting the principled arguments for dismissing the relevance of neuroscience. On the other hand, there was increasing emphasis in neuroscience itself at the *systems* level, where researchers focused on the organization of the visual system or memory systems rather than the cellular or molecular level. A major factor in this move to the systems level was work on mapping the brain. Beginning in the 1960s, neuroanatomists were increasingly able to use patterns of connectivity between cells to infer the existence of specialized areas (those that were highly interconnected). This strategy produced a finer level of resolution than Brodmann's areas, and in many cases the newly identified areas could be analyzed in terms of their contribution to cognitive functions in the way Brodmann had envisaged (see Article 4, BRAIN MAPPING).

To take the best-developed example, the mapping of visual processing pathways, Cragg (1969) and others identified five different pathways from V1 (for visual area 1, located at the rear of the brain in the occipital lobe) to nearby areas labeled V2, V3, superior temporal sulcus, V4, and V4a. By providing various visual stimuli and recording from cells in each of these areas, researchers such as Semir Zeki (1976) discovered that cells in different regions responded to different types of information – e.g., those in V2 to binocular disparity, those in V4 to colors, and those on the superior temporal sulcus to motion (see Article 34, SINGLE NEURON ELECTROPHYSIOLOGY). Then Mortimer Mishkin and Leslie Ungerleider, relying primarily on lesion studies in monkeys, differentiated two main routes for processing visual information, as shown in figure I.11 (Mishkin et al., 1983). One route proceeds ventrally into the posterior temporal cortex along the inferior longitudinal fascilus to areas TEO and TE. Based on the fact that lesions in the posterior temporal area result in loss of pattern discrimination and that lesions in TE in particular result in failure to recognize previously presented objects, Mishkin and Ungerleider proposed that this pathway analyzed the physical properties of visually perceived objects, such as size, color, texture, and shape. One feature of this ventral pathway is that neurons further along the pathway have increasingly large receptive fields, suggesting that the more distant neurons take responsibility for recognizing objects independently of where they appear in the visual field. The other route proceeds dorsally into the posterior parietal cortex. Lesions in the posterior parietal cortex in monkeys produce an inability to select the one of two food wells that is closer to a visual landmark, suggesting that this pathway figures in perception of spatial relations.

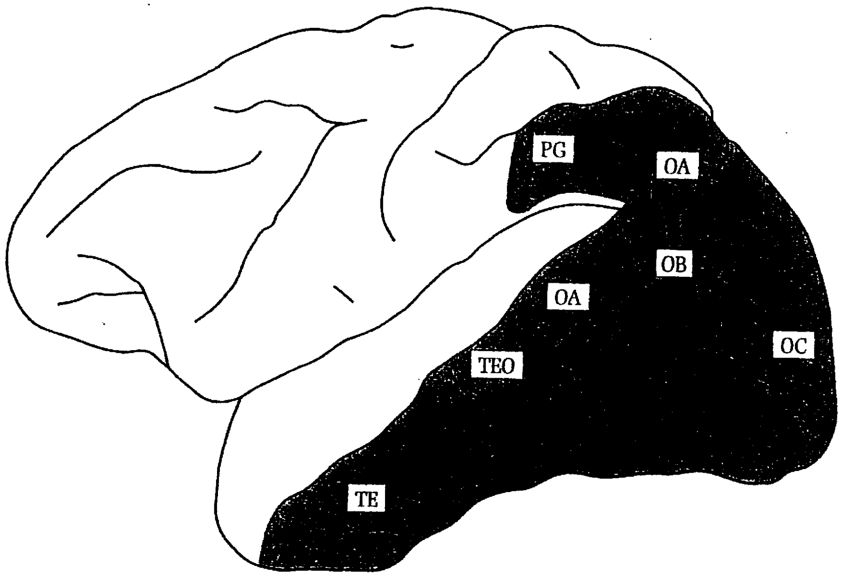

Figure I.11 Two pathways of visual processing in the rhesus monkey proposed by Mishkin, Ungerleider, and Macko (1983). Each begins in area OC (primary visual cortex, also called V1) and projects into prestriate areas OB (V2) and OA (V3, V4, and MT). The *what* pathway then projects ventrally into inferior temporal cortex (areas TEO and TE), whereas the *where* pathway projects dorsally into inferior parietal cortex (area PG).

The two systems differentiated by Mishkin and Ungerleider quickly became known popularly as the *what* and the *where* systems. Research on these pathways has continued. One discovery is that the two pathways diverge even before V1 and employ different cells in V1. Another is that the visual system has a surprisingly large number of different processing areas (Felleman and van Essen, 1991, distinguish 32 areas in the Macaque monkey), and that there are numerous connections between them (Felleman and van Essen identify more than 300). Of particular interest, many of these connections run between the what and the where pathways. Although the story is proving to be ever more complex, it is possible, especially via SINGLE NEURON ELECTROPHYSIOLOGY and NEUROIMAGING, to determine what kinds of information the cells in particular areas are most sensitive to, and to begin to model how information is processed in the visual cortex (van Essen and Anderson, 1990). (For another example of how discoveries of neural architecture are providing guidance in discovering the processing involved in performing cognitive functions, see Article 13, LANGUAGE EVOLUTION AND NEUROMECHANISMS.)

With recognition growing in the 1980s that a variety of research endeavors in neuroscience were relevant to cognitive modeling, neuroscientists and cognitive scientists increased their communication and began to actively collaborate. The name *cognitive neuroscience* has been adopted for this growing enterprise. A number of major researchers on both sides of the intellectual aisle played a key role. We will focus on three: two from the cognitive science side, Steve Kosslyn and Michael Posner, and one from the neuroscience side, Michael Gazzaniga.

From his days as a psychology graduate student at Stanford (see section 2.4), Kosslyn's research focused on the use people make of mental imagery in solving problems. (See Article 12, IMAGERY AND SPATIAL REPRESENTATION.) He soon became embroiled in a vigorous debate as to whether the underlying mental processes operate on representations employing a depictive (quasi-pictorial) format or on ones employing a propositional format. Kosslyn (1980) argued for a depictive format and supported his view with data from a variety of ingenious tasks. He found, for example, that scanning a mental image for a small property (e.g., a cat's claws) took longer than scanning it for a large property (e.g., a cat's head). The case against a depictive format was argued most forcefully by Zenon Pylyshyn (1973, 1981). From his key premise that minds (like computers) store information in structures with a sentential syntax, Pylyshyn concluded that mental representations must be propositional. Examining the available data from experiments, he saw no incompatibility with his claim that subjects perform imagery tasks by operating on propositional representations.

The issue between Pylyshyn and Kosslyn is not whether something that *looks* like a picture, rather than a sentence, is before the mind's eye; it is conceded by propositionalists that mental images may seem introspectively to be pictorial. The issue is whether the actual image, as an underlying representation or mental code, is picture-like or proposition-like. The way pictures represent referents is different in various respects from the way in which propositions or sentences represent them. One of these differences (emphasized by Pylyshyn) is that pictures have spatial properties, whereas propositions do not. Compare, for example, a map which tells how far Paris is from Madrid with a sentence that provides that information. The sentence may simply report the number of miles; from the map we will also be able to tell that Madrid is south of Paris, that there are highways connecting the two cities, and so on.

In 1978, John Anderson contended that the imagery debate could not be resolved by behavioral data, since pictorial accounts could always be mimicked by propositional accounts, and vice versa. Accordingly, in more recent work Kosslyn (1994) and other psychologists, including his former student Martha Farah (1989), have employed a variety of neuroscience tools, including studies of DEFICITS AND PATHOLOGIES and NEUROIMAGING, to show that the same areas of the cortex are involved in both vision and visual imagery, and that these areas use topographical (hence depictive) representations.

Michael Posner's research, like Kosslyn's, began firmly rooted in psychology, and he too gradually came to regard links with neuroscience as crucial. In his early research, Posner was a major developer of chronometric methods for identifying elementary cognitive operations (see Article 27, BEHAVIORAL EXPERIMENTATION). In so doing, he was a modern-day counterpart to the Dutch psychologist Frans Cornelis Donders (1868/1969), who pioneered the use of reaction time patterns to analyze cognitive activities into component operations. Donders developed what is known as the *subtractive method*, in which a researcher has a subject perform two tasks (e.g., discriminate between two signals) which are thought to differ in that one involves an additional mental operation not required for the other (e.g., simply react to a signal). The difference in reaction times yields an estimate of how much time the additional operation requires. An objection to the subtractive technique is that it assumes that the additional processing is a pure insertion of a new elementary operation that does not alter the execution of those operations that are common to the two tasks. Saul Sternberg (1969) extended Donders's method into an *additive factors* approach: the researcher uses multiple manipulations, each of which should have an independent effect on the time required for a particular component operation. Sternberg had subjects memorize a short list of items and then say whether or not a test item was on the memorized list. Time to encode the test item should be affected by the clarity of its visual display, and time to mentally scan the memorized list should be affected by its length. If the slope of the list length effect is the same for both clear and degraded test items, then these factors are additive, and one has evidence of separate processing components. A problematic feature of both the subtractive and the additive factors approaches is that they assume that mental operations are performed sequentially, an assumption that is sometimes false (see Article 53, STAGE THEORIES REFUTED). Partly to overcome this objection, Posner often based his conclusions on interference effects (the extent to which performing an additional activity would increase the time required to perform a primary task). Interference indicates the use of a shared resource, whereas no interference indicates independent operations.

Many of Posner's studies of mental chronometry involved mechanisms of ATTENTION while performing two tasks. In one of Posner's experiments subjects were asked to indicate by a key press with one hand whether two letters were the same, while with the other hand they press a different key whenever they hear a tone. Following a visual warning, the two letters were presented sequentially on a visual display. Reaction times to the tone were not affected during the interval between the warning signal and the first letter, but were lengthened during the interval between the two letter presentations. This indicated that an attentional resource was consumed during the interval between letters but not during the earlier interval. While chronometric techniques work from the behavioral level to identify elementary mental operations, Posner foresaw that they might provide the links to underlying neural systems (Posner and McLeod, 1982, p. 478). In 1982 Posner, Roy Pea, and Bruce Volpe published a prescient paper entitled "Cognitive-neuroscience: developments toward a science of synthesis." They focused on a number of techniques by which one might relate neuroprocesses to the elementary mental operations isolated by means of chronometric methods. These included studies of individuals with brain lesions, changes in electrical potentials recorded on the scalp (event-related potentials or ERPs), and measurements of blood flow by means of *positron emission tomography (PET)*, then under development at Washington University in St Louis (where Marcus Raichle was exploring its potential for studying mental processes).

From 1985 to 1988 Posner left the University of Oregon to collaborate with Raichle and Steven Petersen at Washington University in developing the use of PET as a tool for imaging brain processes during the performance of cognitive activities (see Article 32, NEUROIMAGING). A key to this development was a transformation of Donders's subtractive method from a temporal context to a spatial one. Again, subjects were asked to perform two tasks thought to differ in just one component. Instead of seeking differences in reaction times, these investigators looked for differences in blood flow as indicative of which brain regions were most involved in that component. The collaboration between Posner, Raichle, and Petersen produced a series of influential studies. In a 1988 *Science* paper with Peter T. Fox, they constructed a series of word tasks that involved progressively more cognitive operations and analyzed PET data subtractively in order to identify brain regions that were distinctively active for particular operations. For example, subjects showed more brain activity in an area in the anterior left frontal lobe when they were required to generate and say aloud a verb semantically associated with a visually presented noun than when they only had to read the noun aloud. Even in this research Posner has continued to pursue his interest in mechanisms of attention, and his more recent work integrates chronometric measures, neuroimaging, and ERP results in the attempt to distinguish different attentional mechanisms. (See Posner and Raichle, 1994, which also discusses *functional magnetic resonance imaging (fMRI)*, a new technique for neuroimaging that has advantages and disadvantages compared to PET.)

On the neuroscience side, Michael Gazzaniga was one of the chief initiators of greater collaboration between neuroscience and cognitive science. His early work focused on the study of split-brain patients: that is, patients whose corpus callosum (the major neural pathway between the two hemispheres) was severed. Developed by Rochester, New York, neurosurgeon William Van Wagenen, this procedure was used to relieve symptoms in a small proportion of epileptic patients. Initially it seemed not to result in any cognitive deficits. But as a graduate student and postdoctoral fellow with Roger Sperry at the California Institute of Technology, Gazzaniga demonstrated that when communication between the hemispheres was impaired, patients indeed suffered cognitive deficits. Gazzaniga uncovered these deficits by providing stimuli to sensory receptors that passed information to one hemisphere and requiring responses to be made by motor systems controlled by the other hemisphere. This involved, for example, presenting visual stimuli briefly to the left or right of a fixation point. Information from stimuli presented to the left of the fixation point would reach only the right hemisphere, and split-brain patients in whom language production was lateralized in the left hemisphere would be unable to name them. One of Gazzaniga's major findings came from a split-brain patient, P.S., who had suffered considerable damage to the left hemisphere early in his life. One consequence of early damage to language areas in the left hemisphere is greater involvement of the right hemisphere in language tasks. As a result, P.S. was able initially to respond to a question, a crucial part of which was visually presented to the right hemisphere, by spelling out the answer with Scrabble letters; subsequently, he was able to respond orally as well (Gazzaniga and LeDoux, 1978). At this later stage, if the word *cupcake* was presented so that *cup* was seen by the right hemisphere and *cake* by the left hemisphere, he would report seeing the words *cup* and *cake*, but not *cupcake*, indicating that the two hemispheres had each processed part of the word, but were not able to integrate their results.

In addition to his own attempt to integrate cognitive and neuroscience techniques in analyzing split-brain patients, Gazzaniga became a major force in developing the institutional structures required for cognitive neuroscience. At Dartmouth and, more recently, at the University of California, Davis, Gazzaniga has developed academic programs emphasizing cognitive neuroscience (he recently returned to Dartmouth). Since 1989 he has directed annual summer institutes in cognitive neuroscience which have been designed to bring together researchers with different primary research orientations but sufficient background to enter into a common dialogue. In 1988 he created a new journal, *Journal of Cognitive Neuroscience*, which he continues to edit, and in 1993 he founded the Society for Cognitive Neuroscience. For MIT Press he edited a hefty tome, *The Cognitive Neurosciences*; published in 1995, this volume offers a comprehensive review of work at the interface of neuroscience and cognitive science.

One factor in the renewed interest in the relation of the brain to cognitive science was foundation support. In 1984, in the context of its existing cognitive science initiative, the Sloan Foundation instigated another initiative devoted to linking cognitive science and neuroscience. As with the original cognitive science initiative, this support initially emphasized conferences: "One arrangement that seems especially well suited to fostering productive interaction between these two groups of researchers is indeed simple – regular meetings of working groups in which no more than a dozen scientists come together two or three times a year to discuss research topics and strategies" (report for 1984, p. 20). One of these grants went to Johns Hopkins University for the purpose of supporting a series of meetings between neuroscientists and cognitive scientists concerned with memory:

> The study of memory is a central concern in both neuroscience and cognitive science. In neuroscience, memory is studied by neurophysiologists to determine the synaptic changes that underlie neural plasticity, by molecular biologists to determine the molecular processes governing synaptic behavior, and by neuroanatomists to locate the major brain centers that mediate memory. In cognitive science, memory is studied by computer scientists interested in building electronic learning systems and by cognitive psychologists to understand the performance of human memory. These lines of research have developed independently of each other in the past. (Ibid., pp. 20–1)

The following year this project became refined into a sub-initiative targeting computational neuroscience, an endeavor in which tools of mathematics and computer science are employed to model neural processes (extending from the level of single neurons up to the level of coordinated systems involved in vision or action). In 1986 and 1987, 11 computational neuroscience grants ranging from $100,000 to $300,000 were made to collaborative teams, often from more than one institution. The largest of these went to David Sparks, Frank Amthor, and Michael Friedlander at the University of Alabama, Birmingham, for developing and testing computational models of perception of direction of motion, control of visual attention, and control of saccadic eye movements. Their investigations involved collaboration with theorists at MIT, California Institute of Technology and the National Institutes of Health (NIH).

In 1987 the System Development Corporation sponsored a symposium that led to publication of *Computational Neuroscience*, a large collection of papers which attempted to define the field (Schwartz, 1990). Research in this field has expanded dramatically in recent years. Starting with CNS*92 in 1992 in San Francisco, there has been a Computational and Neural Systems meeting annually. The *Journal of Computational Neuroscience* commenced publication in 1994.

More recently, two other foundations joined forces in supporting the building of bridges between cognitive science and neuroscience: the James S. McDonnell Foundation and the Pew Charitable Trusts. In 1989 they awarded substantial grants to establish centers for cognitive neuroscience at eight institutions: the University of Arizona, UCSD, Dartmouth College (later moved to the University of California, Davis), Johns Hopkins University, MIT, the Montreal Neurological Institute, the University of Oregon, and Oxford University. In addition, they have provided training grants and individual research grants as well as support for the summer institutes directed by Gazzaniga.

3.3 Rediscovering the environment: ecological validity and situated action

In its early period cognitive science tended to limit its focus to events presumed to be taking place within the mind/brain. While all researchers would acknowledge that minds exist within bodies and that these bodies have to deal with the external world (both physical and social), most researchers assumed that they could disregard these considerations when studying cognition. Cognition focused on the processing of information inside the head of the person. In order for this to happen, information had to be represented mentally; cognitive processes could then operate on these representations. Subsequently, represented information had to be translated into commands to the motor system, but this took place after cognitive processing as such was finished. Jerry Fodor (1980) articulated such theoretical justification for ignoring both the external world and the body in cognitive science, labeling the resulting framework *methodological solipsism*, but opposition was already gathering in a number of quarters.

One of the major inspirations for challenging methodological solipsism was the work of J. J. Gibson, a psychologist working at Cornell contemporaneous with the early period of cognitive science but whose impact fell elsewhere. Gibson studied visual perception, but instead of concentrating on the information processing going on within individuals as they see, he examined the information that was available to the organism from its environment. His major contention was that there was much more information available *in the light* than psychologists recognized, and that organisms had only to *pick up* this information (Gibson, 1966). They did not need to construct the visual world through a process of inference or hypothesis formation. He argued, for example, that people do not need to construct a three-dimensional representation of the world; rather, there is information specifying the three-dimensional nature of the visual scene in the gradients of texture density, changes in occlusion of objects as the perceiver moves about in the environment, and so forth. One of Gibson's major contentions was that the perceiver must be understood as an active agent using its own motion to sample information about the environment. Gibson also stressed that not all organisms pick up the same information from the environment, but rather would *resonate with* information that is coordinated with their own potential for action. Accordingly, he introduced the notion of an *affordance*; different objects afford different actions to different agents (e.g., a baseball affords throwing to us, but not to frogs), and it is these affordances which organisms are tuned to pick up. (See Article 18, PERCEPTION.)

Gibson launched what has come to be known as the *ecological* approach to perception, one that has been pursued by a number of researchers who have made important discoveries about the sorts of information available to cognitive agents. A well-known example is Gunnar Johansson's (1973) use of motion pictures of people walking in a dark room with lights attached to their ankles, knees, shoulders, elbows, and wrists; he demonstrated that observers were able to see the form and motion of a walking human in a brief film clip. In another example, from studying plovers (shorebirds) diving into water to catch fish, David Lee (1976) identified a variable τ, defined as the inverse of the relative rate of expansion of the image of the object in the visual field, which predicts the time remaining before impact with the object. Plovers apparently track the value of τ as they are diving and use that to determine the moment to close their wings. Finally, developmental psychologist Eleanor Gibson applied the ecological approach to understanding the development of perceptual competence in infants (e.g., her well-known *visual cliff* experiments) and in older children (e.g., her studies of how children learn to differentiate the critical features in a set of meaningless shapes). She produced a theory of perceptual learning and development that emphasized detection rather than construction of information.

Although Gibson was generally either ignored or severely attacked by cognitive scientists (see Fodor and Pylyshyn, 1981, and Ullman, 1980), his emphasis on the ecological components of perception strongly influenced Ulric Neisser after Neisser joined the Cornell faculty in 1967. In his 1976 book, *Cognition and Reality*, which appeared less than a decade after his *Cognitive Psychology*, Neisser attempted an integration of the Gibsonian approach with information processing. Unlike Gibson, Neisser did not completely reject the idea of internal information processing, but he departed from the highly mechanistic conception which had become common in computational models. Instead, he adopted Bartlett's notion of a schema, a highly structured internal state acquired in the course of experience which partially determines what one perceives and remembers. But, like Gibson, Neisser emphasized the complex information available to the organism and the role of the organism in navigating its environment. He criticized much experimental work in psychology for failing to achieve *ecological validity* and relying too exclusively on artificial laboratory tasks. Such research failed to reveal the types of information that cognizers gain from their environments and the ways in which they explore their environments to gain information. Instead of Fodor's solipsistic cognizer, Neisser advocated the perceptual cycle (figure I.12), in which perceivers apply schemata to information received from the environment, which then leads to exploration in the environment and the pickup of new information, and so forth.

In subsequent work Neisser has continued to emphasize the importance of an ecological perspective, but he has carried it beyond perception to more cognitive domains such as CONCEPTUAL ORGANIZATION and MEMORY. For example, he has emphasized the importance of studying memory as it functions in real life. Thus, he analyzed John

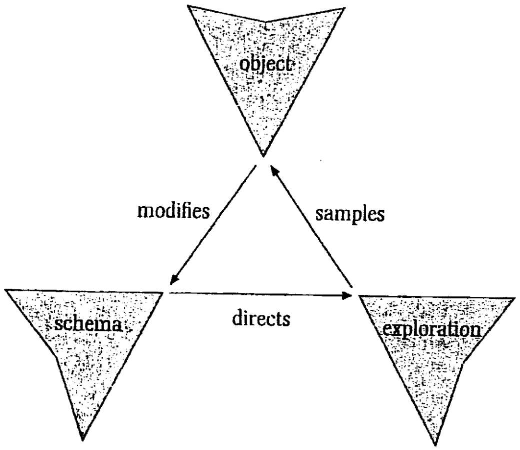

Figure I.12 The perceptual cycle according to Neisser (1975). Perceptual input from an object modifies a schematic representation, which then directs further exploration, leading to further sampling of information from the object.

Dean's testimony on events surrounding Watergate as well as flashbulb memories (memories of when one first learned of major events such as the Challenger crash and the San Francisco earthquake that seem etched in one's mind). He showed in both cases how recall is often erroneous due to the processes by which the subject reconstructs the event. While Neisser attempted to integrate Gibson's perspective with a more traditional cognitive one, others influenced by Gibson such as Michael Turvey, Robert Shaw, and Scott Kelso have pursued a more radical approach, one that is now leading to links with the dynamical systems perspective discussed in section 3.1.

Although the influences stemming from Gibson have been a major factor in the move toward a more ecologically valid cognitive science, there are other major figures who have moved in an ecological direction independently of Gibson. One of these is Donald Norman, whose career has traversed a path from the University of Pennsylvania to Harvard's Center for Cognitive Studies to UCSD (where he helped to establish the LNR group and later the cognitive science program) to Apple Computer to Hewlett-Packard. In his later research he began to focus on how cognizers operate in real world contexts, and especially on ways in which artifacts produced by them favor or impede their performance. (See Article 56, EVERYDAY LIFE ENVIRONMENTS). The focus on how cognition occurs in relation to artifacts in the environment has also been pursued by Norman's UCSD colleague, Ed Hutchins, who has made remarkable studies of how cognition is distributed between agents and artifacts in such activities as ship navigation (see Article 39, EMBODIED, SITUATED, AND DISTRIBUTED COGNITION, and Article 40, MEDIATED ACTION).

A recent development in PERCEPTION has also played a role in reconnecting cognitive systems to their environments. A traditional view of perception, advanced by David Marr among others, is that the task of the perceptual system is to build up a comprehensive representation of the visual scene. But a number of investigators from computer science (Dana Ballard), neuroscience (Terrence Sejnowski), neuropsychology (V. S. Ramachandran), and philosophy (Patricia Churchland) have begun to argue that this badly mischaracterizes what the visual system does. Rather, they argue that it constantly interacts with visual scenes, extracting information relevant to the organism's motoric goals. Because the visual system is capable of rapid eye movements (*saccades*), we can, by using eye movements, gain information about any part of the visual field. This helps generate the illusion that we encode the whole visual field. In addition to eye movements, these researchers emphasize how cognitive agents move their heads and their bodies to gain even more information about the layout of their environments. Churchland, Ramachandran, and Sejnowski (1994), draw heavily upon both behavioral studies of perceptual behavior and findings from neuroscience (in particular, the prevalence of recurrent pathways from higher brain centers down to basic visual areas and the physiological effects of later processing stages on the response patterns of neurons involved in early visual processing).

The recognition that cognitive systems are in constant interaction with their environment has resulted in an infusion into cognitive science of research and methodologies in such social sciences as anthropology and sociology (see Article 30, ETHNOMETHODOLOGY). There has been tension between mainstream cognitive scientists (whose tools are primarily directed at understanding processes inside the head) and those who have embraced social science perspectives (whose tools are concerned with the social contexts of such processes); see, for example, the conflicts in the 1993 special issue of *Cognitive Science* devoted to situated action. However, constructive interactions have increasingly been achieved. One consequence has been the reintroduction of work from the Soviet tradition in psychology that developed from the research of Lev Vygotsky (see Article 40, MEDIATED ACTION).

3.4 *Rediscovering function: cognitive linguistics*

During the middle period in the history of cognitive science, Chomsky's radical new approach had gradually become dominant in linguistics and influential in other disciplines. This is not to say, however, that Chomsky continued to affirm the same grammatical models as he had advanced in the 1950s and 1960s. In his initial generative grammars the emphasis was on transformational rules that could be used to derive surface structures from underlying deep structures. Increasingly, Chomsky has deemphasized transformational rules and the distinction between deep structure and surface structure. In X-bar theory, government and binding theory, and most recently minimalism, the transformational rules have been reduced to the point that minimalism posits just one rule, *move alpha* (move any category anywhere). It generates far more structures than the old transformational rules, but constraints filter out the ungrammatical ones. Partly as a result of the different grammars that Chomsky has advanced over the decades, which some researchers have continued to pursue even after Chomsky himself has abandoned them, partly due to splits from Chomsky and alternative theories within linguistics, and partly due to investigators who have developed grammars for different purposes (such as natural language processing in AI), there is now a rich variety of approaches in LINGUISTIC THEORY. *Optimality theory* (Prince and Smolensky, in press) is particularly distinctive and important. More computational and statistical than Chomskian linguistics, it retains the separation between such components as semantics, syntax, and phonology and has been especially influential in the 1990s in phonology.

A number of linguists led by John R. Ross (a former student of Chomsky), George Lakoff, Paul Postal, and James McCawley broke from Chomsky as early as the 1960s over his insistence on the autonomy of syntax. Chomsky held that syntactic principles are not the product of other linguistic or cognitive processes, hence that one can characterize syntax independently of other aspects of language (e.g., its semantics or the pragmatics involved in its use). Adopting the label *generative semantics*, these early critics sought to extend Chomsky's generative program into the domain of semantics by developing rules that would generate syntactic structures from semantic representations without using a privileged intermediate level, deep structure, to segregate semantic from syntactic parts of the derivation. Chomsky rejected these extensions of his endeavors, and the disputes between these generative semanticists and Chomsky resulted in the *linguistic wars* of the late 1960s and early 1970s (Harris, 1993). After a period of heated controversy, generative semantics morphed into a less cohesive but innovative variety of alternative approaches. Chomskian linguistics retained its cohesiveness but also underwent considerable change.

One generative semanticist, George Lakoff, began in the 1980s at UC–Berkeley to develop a new approach eventually known as *cognitive linguistics*. (There is now a professional society and a journal that employ that name.) Lakoff especially emphasized the structure of concepts and (with philosopher Mark Johnson) did influential work on metaphor. At UCSD, Ronald Langacker and his former student, Gilles Fauconnier, developed a highly systematized theory of cognitive grammar that emphasized the grounding of language in highly abstract spatial representations. In general, cognitive linguists shared a conviction that syntax, far from being autonomous, was the product of more basic cognitive operations. Loosely affiliated with cognitive linguistics are other linguists and psycholinguists who focus more on the pragmatics or function of language use and who try to account for syntactic structures in terms of such functions (see Article 37, COGNITIVE LINGUISTICS, and Article 31, FUNCTIONAL ANALYSIS).

4 Coming of age: downwards and outwards

We have explored the origins and early development of cognitive science, but where is it going? Just as it is hard to know how adolescents will resolve their identity crises and choose from the many paths that lie before them, it would be presumptuous for us to predict the precise form that cognitive science will take as it confronts the choices posed in the previous section. The one thing that seems certain is change: cognitive science is being pulled vertically *down* into the brain and horizontally *out* into the environment.

Where might these lines of growth take the field? As a starting point, let us return to the characterization of cognitive science that we put forward at the beginning:

> Cognitive science is the multidisciplinary scientific study of cognition and its role in intelligent agency. It examines what cognition is, what it does, and how it works.

The changes will involve multiple disciplines; how they develop and interact will determine the shape of cognitive science as we enter the next millennium. We will return to this topic shortly, but first we consider answers to the questions of what cognition is, what it does, and how it works.

The cognitive science of the 1970s answered these questions from an information processing perspective. *What cognition is*: the processing of information in the head. *What cognition does*: it enables an agent to exhibit intelligent behavior, which is prototypically manifested in such activities as solving the Tower of Hanoi problem or understanding sentences. Psychologists confined such activities to the laboratory for study, and AI researchers modeled them in programs which might be judged adequate to the extent that they passed the Turing test (by generating behavior indistinguishable from that of a human). *How cognition works*: like a computer. Information is encoded in a symbolic representational format upon which rules operate – much like algorithms in a programming language.

Through the identity crisis that we explored in section 3, cognitive scientists began to reassess some of these answers. We can look at the new thinking on the three questions in reverse order. *How cognition works*: it proved difficult to accommodate some of the data about how humans actually behave without going beyond the computer metaphor. Even rich data structures like schemas and frames were inadequate to capture the fluid character of human cognition. To some cognitive scientists the neurally inspired approach of connectionism offered a way to realize previously neglected characteristics of human intelligent activity, such as graceful degradation and soft constraint satisfaction. The return to the brain, though, has brought more than a new computational framework. Minimally, many cognitive scientists would insist that any answer to how cognition works must be compatible with emerging knowledge of how the brain works. Others go much further, maintaining that the best way to gain clues as to how cognition works is to study the brain. This does not mean giving up a computational perspective, since many researchers in cognitive neuroscience take as their focus developing computational models of neural activity. It does mean, however, that the inspiration for developing accounts of how cognition works is no longer the digital computer; instead, knowledge about how the brain works increasingly provides the foundation for theoretical modeling.

Another element in the identity crisis involved reassessment of *what cognition does*. The return to the environment and the body refocused attention on how cognition facilitates life in the real world. Although many cognitive psychologists have continued to emphasize laboratory studies of problem solving, reasoning, memory, and language processing, the new concern for *ecological validity* has redirected others to study skills and abilities as exercised in the real world. These have included the old abilities studied in new contexts (a server in a restaurant remembering orders), as well as abilities newly of interest (a navigator guiding a large ship into a harbor). In AI, the Turing test began to lose its status as a sufficient indicator of intelligent agency in artificial systems, and increasing numbers of researchers turned from modeling rational thought to such projects as building robots that could operate in real environments. It became recognized that what cognition does is to provide an open-ended capacity to respond appropriately and flexibly to whatever may come along or appear. It provides a capacity to respond reasonably, as Descartes remarked in the *Discourse on Method*, to "the contingencies of life."

Neither of these developments led to a reassessment of *what cognition is* – the processing of information within the head of the agent. Very recently, however, even this answer has begun to be questioned. We consider three aspects of the challenge in turn.

(i) On one front, investigators emphasizing MEDIATED ACTION and EMBODIED, SITUATED, AND DISTRIBUTED COGNITION are questioning whether analyses that separate and focus on activities *within* the brain are adequate. The alternative view is that interactions between the brain and the environment within which it functions are so intricate and pervasive that the primary unit of analysis must be the system formed by the interacting brain and the environment.

(ii) Expanding the boundaries of the cognitive system so as to incorporate parts of the world is supported by some of the advocates of a move beyond connectionism to dynamical systems theory. In dynamical models, researchers seek to identify a variety of parameters that affect the performance of a system and to develop mathematical laws, frequently in the form of equations employing first or second derivatives, that describe the changes in the system over time. Such mathematical accounts do not impose a boundary at the skin. Even if one does develop a dynamical model limited to processes occurring inside the skin, the fact that one can always couple dynamical systems that share parameters into a compound system ensures the potential for linking models of activities within the brain to those in the environment in a single theoretical model (see Article 38, CONNECTIONISM, ARTIFICIAL LIFE, AND DYNAMICAL SYSTEMS).

(iii) On a third front, as researchers attend more to the brain, their attention is drawn to aspects of mental life, such as EMOTIONS and CONSCIOUSNESS, which may not be best described in terms of information processing, at least as it has been understood so far. Emotional responses are largely under the control of midbrain structures that comprise the limbic system, rather than cortical structures, and these systems do not seem to work by encoding and processing information. While it has been acknowledged that emotions and consciousness modulate cognitive activity, cognitive scientists have frequently assumed that they could disregard their impact and study cognition in isolation. But some cognitive scientists are coming to believe that responsiveness to the contingencies of life relies in significant part on these other responses of the brain. A domain such as chess is highly restricted and may be mastered by computers employing sophisticated inference strategies. A rook is just a rook. But let our eyes sweep across a great landscape, and the richness of sensory detail seems informationally overwhelming. This is because, claim some cognitive scientists, visually experienced detail *is* beyond information processing. The conscious and emotional world has a character, a subjective quality, an identity, which cannot be captured on a purely information processing account. Meanwhile, that subjective quality helps us to respond appropriately to life's contingencies (see Lahav, 1993).

A comprehensive answer to the question of what cognition is, if it is not limited to information processing within the brain, has not yet been developed. Accordingly, it is not yet clear whether information processing will turn out to be a component of a more comprehensive characterization of cognition, or whether it will have been a false step. Given that what cognition does is to enable agents to interact intelligently with their environments, it seems plausible that in some way information processing will be a part of a more adequate conception of cognition. However, the recent developments we have touched upon suggest that it will not exhaust cognition.

Given these views about what cognition is, what it does, and how it works, let us return to the question of what disciplines figure in cognitive science's attempt to understand it. In the phase of gestation, the three principal disciplines from which researchers began to interact and formulate a plan of study were neuroscience, psychology

(especially cognitive), and computer science (AI). Drawing from sources such as cybernetics and information theory, researchers from all three disciplines proposed that cognition should be understood computationally, with computational models being constrained by neural and behavioral data. This interaction is represented graphically as a triangle in figure I.13a, with the three disciplines represented by the three vertices. The fact that the names of the disciplines are all in the same font size indicates that they were roughly equal contributors to the new enterprise. Following the format of figure I.10 above, the lines connecting the nodes represent the interdisciplinary interactions. The fact that they are of the same width indicates that the interactions were equally potent.

Around the crucial year 1956, however, major changes in the collaborations occurred. Figure I.13b illustrates the new mix of fields and interdisciplinary connections that characterized cognitive science during the period of its initial maturation (roughly until 1985, as discussed in section 2). A new program in linguistics, advanced by Noam Chomsky, boosted its influence to about the same level as psychology. And at nearly the same moment, neuroscience began to play a much less significant role. Three other disciplines – sociology, anthropology, and philosophy – began to play an ancillary role. Finally, computer science (AI) became the preeminent cognitive science discipline during this period of initial maturation.

As in figure I.13a, relative prominence and influence are indicated by font size, but with the additional disciplines, the initial triangle has become a hexagon, and many more interdisciplinary interactions are possible. Some of these are better developed than others, and the thickness of the lines connecting the various disciplines are intended to be indicative of the extent to which these connections were pursued during the period of initial maturation.

Recently, the relative importance of the various contributing disciplines has changed yet again, as illustrated in figure I.13c. As we have noted, neuroscience is playing an increasingly important role in cognitive science. Techniques of functional investigation in neuroscience are becoming increasingly important in guiding thinking about cognition. Social and cultural studies originating in sociology and anthropology are also coming to a play a more influential role, whereas both computer science (AI) and linguistics have become less influential. The diminished role of computer science (AI) is due to at least two factors. On the one hand, internal pressures in computer science to focus on the development of such basic computer tools as computer operating systems and compilers have provided a less hospitable home for AI. On the other hand, computational modeling tools, once primarily the property of computer scientists interested in artificial intelligence, are becoming tools of practitioners in other cognitive science disciplines. Thus, these disciplines are all represented in the same size font as philosophy, whose importance remains stable.

As important as the transitions in the significance of individual disciplines, however, has been the increased importance of some of the interdisciplinary interactions. One thriving area of inquiry is cognitive neuroscience, which combines the behavioral tools of cognitive psychology with the functional methods of neuroscience in identifying the brain areas involved in performing cognitive tasks. The field became established as different sources of information about the contribution of different brain areas to cognition became available: first deficits and pathologies, then single neuron electrophysiology, then event-related potentials in electroencephalograms. It is thriving

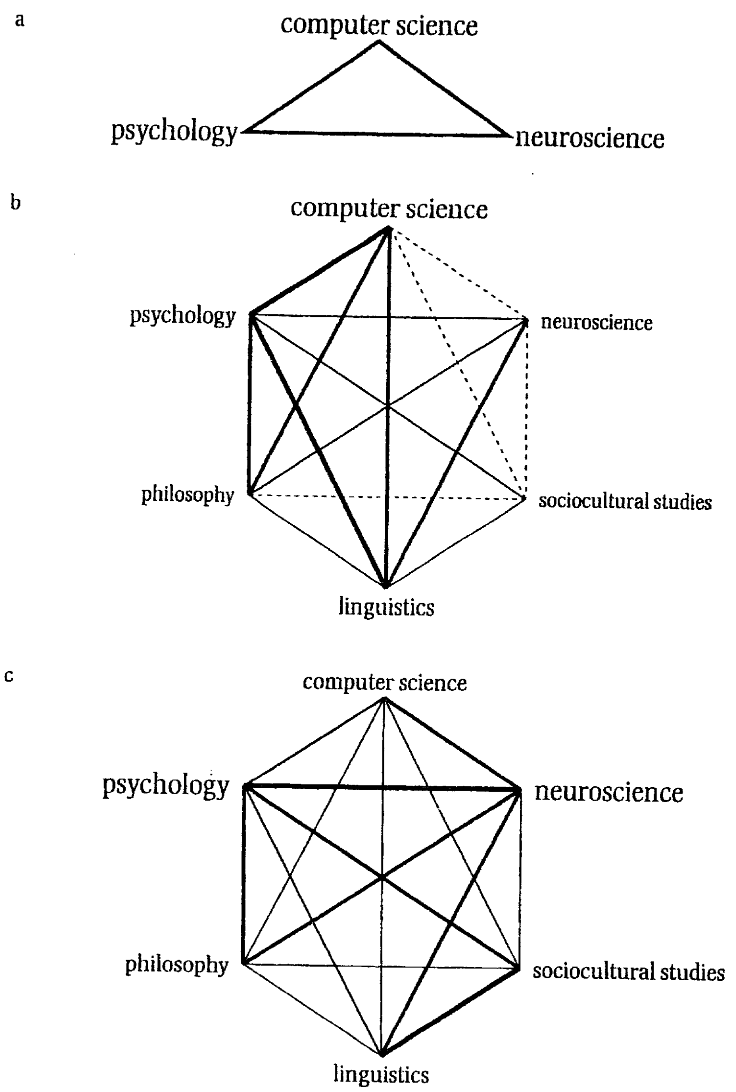

Figure I.13 Contributing disciplines and interdisciplinary connections during three different stages in the development of cognitive science. The font size in which the name of the discipline is printed reflects the relative importance of that discipline to cognitive science during the stage in question. The lines between disciplines represent interdisciplinary connections. The thickness of the line represents the activity level of that interdisciplinary connection; a dotted line represents an essentially undeveloped connection. a. The three disciplines that were central to cognitive science during most of its gestation (linguistics became an important contributor only at the very end of this period). b. The six disciplines contributing to cognitive science during the period of initial maturation. c. The six disciplines currently contributing to cognitive science.

now, due to the advent of neuroimaging, which provides relatively direct information from intact human brains at good temporal and spatial resolutions. Also newly important is the impact of sociology and anthropology on both cognitive psychology and computer science (AI), as evidenced in the role of mediated and situated action in some psychological models and robotic simulations.

Some of the already established interdisciplinary connections have changed their character. For example, the interaction between psychology and linguistics took the form of psycholinguistic inquiries into the psychological reality of grammar during cognitive science's initial maturation. Interactions are more pluralistic in the current era; for example, cognitive linguistics denies the autonomy of syntax and instead appeals to cognitive processes to try to explain the link between meanings and phonological forms. Optimality theory can retain an autonomous syntax but proposes to account for syntax in terms of a system of soft constraints. Meanwhile, Chomskian language development researchers study parameter setting within an autonomous syntax module that retains a classic architecture.

Figure I.13c provides a picture of the disciplines of cognitive science and their interrelations that are likely to figure in the immediate future. But this picture is no more likely to remain static than the previous ones. At different times different disciplines will be better positioned than others to advance our understanding of cognition. Indeed, we can identify one field which is just beginning to make contributions and may soon become a highly important player: behavioral psychopharmacology (see Willner, 1991). For the most part, researchers have focused on electrical properties of the brain in seeking to understand its relation to cognition. But a fundamental discovery of early twentieth-century neuroscience was that the nervous system is not reticular: neurons are separated by synapses. Communication across synapses is mediated by chemicals. Starting from a few key neurotransmitters, researchers have identified a large class of chemicals that serve this function. These have been shown to be critical to normal cognition, but a detailed understanding of how they figure in cognition remains to be developed.

If the nature of chemical processes in the brain turns out to be critical to cognition, this will certainly strengthen the tendency in cognitive science to pursue a downwards direction of inquiry. But the development of behavioral psychopharmacology may, as the name suggests, also direct researchers outwards to the environment. The effectiveness of some psychopharmacological agents is linked with, or dependent upon, changes in the environment of subjects using the particular drug (see Whybrow, 1996). Behavioral psychopharmacology may thus encourage further efforts by cognitive scientists both in searching for underlying mechanisms and in examining relations to environments.

Thus, cognitive science has been, and promises to remain, broadly interdisciplinary. Interdisciplinary research always involves a tension, since practitioners of different disciplines generally bring to the interaction different agendas, different research tools, and different models of satisfactory answers. Successful interdisciplinary research requires rendering these differences compatible. Sometimes, despite the loftiest goals, interfield collaboration founders. Cognitive science has experienced its share of tensions. Conflicts between symbolic modeling (inspired by the digital computer) and neural network modeling (inspired by the brain) are a current case in point. Another has been the conflict over the autonomy of syntax advocated by Chomskian linguistics and rejected by the cognitive and functional grammarians. So far, the institutions of cognitive science, such as the journal *Cognitive Science* and the Cognitive Science Society, have proved capable of remaining fairly inclusive.

But will cognitive science remain viable as a field of interdisciplinary cooperation? The simultaneous pulls *downwards* into the brain and *outwards* into the world may prove to be too much pulling, and lead to the disintegration of cognitive science. On the other hand, awareness that brain processes and events in the external world interact in crucial ways may suffice to hold the inquiry together. The attempt to combine the two in this volume represents the conviction that, to at least a significant degree, cognitive science will continue as a robust interdisciplinary endeavor.

There is a feature of the competing pulls downwards and outwards, though, that increases the risk of a serious rift. One important aspect of the information processing perspective that was adopted in cognitive science was the attempt to specify mechanisms underlying cognition. Each operation upon information represented a process occurring within the cognitive system. Behaviorists and mathematical psychologists had not attempted to identify such mechanisms; thus, the cognitivists' rebellion against behaviorism and mathematical models of learning involved embracing a different conception of explanation – one in which it was not sufficient to identify laws or mathematical regularities in behavior, but actual mechanisms responsible for it. The current turn downwards into the brain is sometimes represented as *reductionistic*. Since the word *reduction*, however, is understood in a host of different ways, we will not employ it here. What it clearly represents is a further step in the continuing quest to identify the underlying mechanisms responsible for intelligent behavior.

In treating the quest for explanation as a quest for mechanisms, cognitive scientists are adopting a perspective on explanation that has been widely shared in the life sciences. For example, biologists seek to explain such processes as energy liberation and reproduction in animals by characterizing the mechanisms which make these possible. Discovery of these mechanisms has generally involved identifying a function that a system performs (providing energy for work, or comprehending and producing linguistic utterances), *decomposing* that function into component functions, and *localizing* those component functions in the physical system (Bechtel and Richardson, 1993). Sometimes the decomposition may produce a linear sequence of component functions, but it need not (see Article 53, STAGE THEORIES REFUTED). The performance of different tasks may be highly integrated (e.g., through backwards or recurrent connections). Moreover, sometimes localization will identify one discrete area of the system responsible for a component task, but other times the component function may be distributed throughout the system. Further, while researchers may aspire ultimately to identify the actual physical locus, often they must settle for indirect evidence that such a locus exists (e.g., by demonstrating that each of two different functions can be preserved while the other is incapacitated: a double dissociation). Cognitive science, especially cognitive psychology, can be seen as proposing decompositions of the cognitive system. Without invoking neuroscience, the evidence for the underlying mechanisms performing the different functions remains indirect. The support for these mechanisms increases when one combines behavioral and neural sources of evidence.

While the downward pull has not challenged the emphasis on mechanism in cognitive science, some researchers pursuing the outward pull have questioned it and have advocated a return to an explanatory framework that in some respects resembles that of behaviorists and mathematical psychologists. Thus, some advocates of dynamical analyses (van Gelder, 1995) suggest that it is sufficient to identify critical variables characterizing the state of systems and to construct mathematical laws to account for the ways in which the values of these variables change over time. These theorists reject the idea that information is represented in the system, and that it is representations that are operated on by different components of the system (see Article 50, REPRESENTATION AND COMPUTATION). That is, they reject the information processing perspective.

If there is incommensurability between mechanistic and dynamical models, and if it tends to correspond to the difference between efforts to go downwards into the brain and outwards into the environment, then we may have the seeds of a significant fracture in cognitive science. The developmental psychologist Erik Erikson (1968, p. 136) coined the appositely awkward expression *distantiation* to identify what happens in the life of an individual who repudiates elements in her or his personality which seem incompatible or threatening. As the areas of cognitive neuroscience and dynamical systems theory are progressively delineated, they may fortify their gains by overvaluing their differences and hence increasing the distance.

A distantiated future will not satisfy many committed to cognitive science. But those committed to an integrated cognitive science may discover that the potential for fracture is not as serious as it seems. At present the dynamicists' challenge is not fully formed. Central to the challenge are the notions of information processing and representation, but these notions are currently vague and must be theoretically regimented. It may well be that a mature dynamical account will posit genuine information processing and representation, although the representations employed will not be syntactically structured or sentence-like. The model of syntactically structured or sentence-like representations (Fodor's language of thought) is, in any case, under severe attack from a number of quarters in contemporary cognitive science. Other models of representations have come to the fore: graphs, maps, holograms, house plans, and other nonsentential schemes, and many investigators are exploring the idea that the brain may process information using one or more of these other kinds of representation. The path has already been cleared by theorists of perception (see Article 18, PERCEPTION, also Article 19, PERCEPTION: COLOR), who have advocated understanding perception in terms of transformations of spatial or quasi-pictorial representations.

There is further reason to doubt the seriousness of the dynamicist/mechanist rift. On the one hand, neuroscientists often emphasize that recurrent or backward projections in the brain outnumber forward connections, suggesting that the brain itself is a highly interconnected system. Accordingly, they expect that a computation carried out by any given brain part may be highly influenced by activities elsewhere in the brain. In such a complex system, those attempting to develop mechanistic accounts may require the dynamicists' tools. On the other hand, successful dynamical accounts relating different brain and environmental parameters themselves call out for explanation. What underlying mechanisms produce the behavior described at an abstract level by the dynamical equations? Unless one accepts action at a distance, one is back to the search for intermediate processes (in the head), and it is natural to characterize these processes in terms of how they carry information and represent self and world. Accordingly, it may be possible for dynamicists and mechanists to coexist within cognitive science, and even to collaborate with each other.

In the life of cognitive science time is of the essence. Cognitive science is not a static entity. Nor are the disciplines that comprise it. In its gestation period it lacked institutional organization; integration resulted from the power of the idea of information processing that suggested a way to view the activities of the nervous system and release psychology from behaviorism. Between 1960 and 1985 it matured, developing its identity in terms of computational models of a variety of cognitive activities. With the support of generous benefactors, it developed its institutional base. But having reached adulthood, it now recognizes some of the advantages of approaches overlooked in its development and has been drawn back downwards into the brain and outwards into the world. As it pursues its adult career, these and other factors will create tensions. But we are optimistic that cognitive science will not only endure but will develop into an even more interesting domain of science as it confronts challenges not yet recognized, devises theories not yet anticipated, and encompasses stances and deploys methods not yet imagined.

5 References and recommended reading

Alfred P. Sloan Foundation 1975–87: *Annual Report*. New York: Alfred P. Sloan Foundation.

*Anderson, J. A. and Rosenfeld, E. 1988: *Neurocomputing: Foundations of Research*. Cambridge, Mass.: MIT Press.

Anderson, J. R. 1978: Arguments concerning representations for mental imagery. *Psychological Review*, 85, 249–77.

—— 1983: *The Architecture of Cognition*. Cambridge, Mass.: Harvard University Press.

Anderson, J. R. and Bower, G. H. 1973: *Human Associative Memory*. New York: Winston.

Ashby, W. R. 1952: *Design for a Brain*. New York: John Wiley and Sons.

Atkinson, R. C. and Shiffrin, R. M. 1968: Human memory: a proposed system and its control processes. In K. W. Spence and J. T. Spence (eds), *The Psychology of Learning and Motivation: Advances in Research and Theory*, vol. 2, New York: Academic Press, 89–195.

*Baars, B. J. 1986: *The Cognitive Revolution in Psychology*. New York: Guilford Press.

Baddeley, A. D. 1976: *The Psychology of Memory*. New York: Basic Books.

Bartlett, F. C. 1932: *Remembering*. Cambridge: Cambridge University Press.

*Bechtel, W. 1988: *Philosophy of Mind: An Overview for Cognitive Science*. Hillsdale, NJ: Erlbaum.

Bechtel, W. and Richardson, R. C. 1993: *Discovering Complexity: Decomposition and Localization as Scientific Research Strategies*. Princeton: Princeton University Press.

Block, N. 1978: Troubles with functionalism. In C. W. Savage (ed.), *Perception and Cognition. Issues in the Foundations of Psychology*. Minnesota Studies in the Philosophy of Science, vol. 9, Minneapolis: University of Minnesota Press, 261–325.

*Blumenthal, A. L. 1970: *Language and Psychology: Historical Aspects of Psycholinguistics*. New York: Wiley.

Bobrow, D. G. and Collins, A. (eds) 1975: *Representation and Understanding: Studies in Cognitive Science*. New York: Academic Press.

Bradley, D., Garrett, M. and Zurif, E. 1980: Syntactic deficits in Broca's aphasia. In D. Caplan (ed.), *Biological Studies of Mental Processes*, Cambridge, Mass.: MIT Press, 269–86.

Bransford, J. D. and Franks, J. J. 1971: The abstraction of linguistic ideas. *Cognitive Psychology*, 2, 331–50.

Bransford, J. D., Barclay, J. R. and Franks, J. J. 1972: Sentence memory: a constructive versus an interpretive approach. *Cognitive Psychology*, 3, 193–209.

Broadbent, D. E. 1958: *Perception and Communication*. New York: Pergamon Press.

Broca, P. 1861: Remarques sur le siège de la faculté suivies d'une observation d'aphéme. *Bulletin de la Société Anatomique de Paris*, 6, 343–57. Trans. by G. von Bonin in *idem* (ed.), *Some Papers on the Cerebral Cortex*, Springfield, Ill.: Charles C. Thomas, 49–72.

Brodmann, K. 1909: *Vergleichende Lokalisationslehre der Grosshirnrinde*. Leipzig: J. A. Barth. English translation by L. J. Garey (1994), *Brodmann's "Localisation in the Cerebral Cortex."* London: Smith-Gordon.

*Bruner, J. S. 1983: *In Search of Mind: Essays in Autobiography*. New York: Harper and Row.

Bruner, J. S. and Goodman, C. 1947: Value and need as organizing factors in perception. *Journal of Abnormal and Social Psychology*, 42, 33–44.

Bruner, J. S. and Postman, L. 1949: On the perception of incongruity: a paradigm. *Journal of Personality*, 18, 206–23.

Bruner, J. S., Goodnow, J. and Austin, G. 1956: *A Study of Thinking*. New York: Wiley.

Bry, A. 1975: *A Primer of Behavioral Psychology*. New York: Mentor/New American Library.

Chomsky, N. 1957: *Syntactic Structures*. The Hague: Mouton.

—— 1959: Review of Skinner's *Verbal Behavior*. *Language*, 35, 26–58.

—— 1965: *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.

—— 1968: *Language and Mind*, 1st edn. New York: Harcourt.

—— 1975: *The Logical Structure of Linguistic Theory*. New York: Plenum.

Chomsky, N. and Miller, G. A. 1958: Finite-state languages. *Information and Control*, 1, 91–112.

Churchland, P. S. 1986: *Neurophilosophy*. Cambridge, Mass.: MIT Press.

Churchland, P. S., Ramachandran, V. S. and Sejnowski, T. 1994: A critique of pure vision. In C. Koch and J. L. Davis (eds), *Large-scale Neuronal Theories of the Brain*, Cambridge, Mass.: MIT Press, 23–61.

Collins, A. 1977: Why cognitive science. *Cognitive Science*, 1, 1–2.

Collins, A. and Quillian, M. R. 1969: Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 8, 240–7.

Cooper, L. A. 1975: Mental rotation of random two-dimensional shapes. *Cognitive Psychology*, 7, 20–43.

*Corsi, P. 1991: *The Enchanted Loom. Chapters in the History of Neuroscience*. New York: Oxford University Press.

*Corsini, R. J. 1994: *Encyclopedia of Psychology*. New York: Wiley.

Cragg, B. G. 1969: The topography of the afferent projections in the circumstriate visual cortex (C. V. C.) of the monkey studied by the Nauta method. *Vision Research*, 9, 733–47.

*Crevier, D. 1993: *AI: The Tumultuous History of the Search for Artificial Intelligence*. New York: Basic Books.

Donders, F. C. 1868: Over de snelheid van psychische processen. Onderzoekingen gedaan in het Psyiologish Laboratorium der Utrechtsche Hoogeschool: 1868–1869. *Tweede Reeks*, II, 92–120. English translation by W. G. Koster (1969), On the speed of mental processes. *Acta Psychologica*, 30 (1969), 412–31.

Dreyfus, H. 1972/1979: *What Computers Can't Do*. New York: Harper and Row.

Ebbinghaus, H. 1885/1913: *Memory: A Contribution to Experimental Psychology*. New York: Columbia Teachers' College.

Erikson, E. H. 1968: *Identity, Youth and Crisis*. New York: W. W. Norton.

Farah, M. J. 1989: Mechanisms of imagery–perception interaction. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 203–11.

Feigenbaum, E. A., Buchanan, B. G. and Lederberg, J. 1971: On generality and problem solving: a case study using the DENDRAL program. In B. Melzer and D. Michie (eds), *Machine Intelligence*, vol. 6, Edinburgh: Edinburgh University Press, 165–90.

Felleman, D. J. and van Essen, D. C. 1991: Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1, 1–47.

Festinger, L. 1957: *A Theory of Cognitive Dissonance*. Evanston, Ill.: Row, Peterson.

*Finger, S. (1994). *Origins of Neuroscience*. New York: Oxford University Press.

Fodor, J. A. 1975: *The Language of Thought*. Cambridge, Mass.: MIT Press/Bradford.

—— 1980: Methodological solipsism considered as a research strategy in cognitive psychology. *Behavioral and Brain Sciences*, 3, 63–109.

—— 1987: *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, Mass.: MIT Press.

Fodor, J. A. and Pylyshyn, Z. W. 1981: How direct is visual perception: some reflections on Gibson's "ecological approach." *Cognition*, 9, 139–96.

—— —— 1988: Connectionism and cognitive architecture: a critical analysis. *Cognition*, 28, 3–71.

*Fodor, J. A., Bever, T. G. and Garrett, M. F. 1974: *The Psychology of Language*. New York: McGraw-Hill.

Fritsch, G. and Hitzig, E. 1870/1960: On the electrical excitability of the cerebrum. In G. von Bonin (ed.), *Some Papers on the Cerebral Cortex*, Springfield, Ill.: Charles C. Thomas, 73–96.

Galanter, E. 1988: Writing *Plans* . . . In W. Hirst (ed.), *The Making of Cognitive Science: Essays in Honor of George A. Miller*, New York: Cambridge University Press, 36–44.

*Gardner, H. 1985: *The Mind's New Science*. New York: Basic Books.

Garner, W. R. 1988: The contribution of information theory to psychology. In W. Hirst (ed.), *The Making of Cognitive Science: Essays in Honor of George A. Miller*, New York: Cambridge University Press, 19–35.

Gazzaniga, M. S. (ed.) 1995: *The Cognitive Neurosciences*. Cambridge, Mass.: MIT Press.

Gazzaniga, M. S. and LeDoux, J. E. 1978: *The Integrated Mind*. New York: Plenum.

Geschwind, N. 1974: *Some Selected Papers on Language and Brain*. Dordrecht: Reidel.

Gibson, J. J. 1966: *The Senses Considered as Perceptual Systems*. Boston: Houghton-Mifflin.

Green, D. M. and Swets, J. A. 1966: *Signal Detection Theory and Psychophysics*. New York: Wiley.

Harnad, S. 1990: The symbol grounding problem. *Physica D*, 42, 335–46.

*Harris, R. A. 1993: *The Linguistics Wars*. New York: Oxford University Press.

Harvard University 1961–9: *Annual Report of the Center for Cognitive Studies*. Cambridge, Mass.: Harvard University Press.

*Hearst, E. 1979: *The First Century of Experimental Psychology*. Hillsdale, NJ: Erlbaum.

Hebb, D. O. 1949: *The Organization of Behavior*. New York: Wiley.

Heims, S. J. 1980: *John von Neumann and Norbert Wiener: From Mathematics to the Technologies of Life and Death*. Cambridge, Mass.: MIT Press.

Hempel, C. 1966: *Philosophy of Natural Science*. Englewood Cliffs, NJ: Prentice-Hall.

*Hilgard, E. R. 1987: *Psychology in America: A Historical Survey*. New York: Harcourt Brace Jovanovich.

Hinton, G. and Anderson, J. A. (eds) 1981: *Parallel Models of Associative Memory*. Hillsdale, NJ: Erlbaum.

*Hirst, W. (ed.) 1988: *The Making of Cognitive Science: Essays in Honor of George A. Miller*. New York: Cambridge University Press.

Hopfield, J. J. 1982: Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79, 2554–8.

Hubel, D. H. and Wiesel, T. N. 1962: Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, 160, 106–54.

James, W. 1890: *Principles of Psychology*, 2 vols. New York: Holt.

Jameson, D. and Hurvich, L. M. 1968: Opponent-response functions related to measured cone photopigments. *Journal of the Optical Society of America*, 58, 429–30.

Jeffress, L. A. (ed.) 1951: *Cerebral Mechanisms in Behavior: The Hixon Symposium*. New York: Wiley.

Johansson, G. 1973: Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, 14, 201–11.

Joshi, A. K. and Hopely, P. 1997: A parser from antiquity. *Natural Language Engineering*, 2, 291–4.

*Kessen, W., Ortony, A. and Craik, F. 1991: *Memories, Thoughts, and Emotions: Essays in Honor of George Mandler*. Hillsdale, NJ: Erlbaum.

Kintsch, W., Miller, J. R. and Polson, P. G. 1984: *Method and Tactics in Cognitive Science*. Hillsdale, NJ: Erlbaum.

Kosslyn, S. M. 1980: *Image and Mind*. Cambridge, Mass.: Harvard University Press.

—— 1994: *Image and Brain: The Resolution of the Imagery Debate*. Cambridge, Mass.: MIT Press.

Kuhn, T. 1962/1970: *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

*Lachman, R., Lachman, J. L. and Butterfield, E. C. 1979: *Cognitive Psychology and Information Processing: An Introduction*. Hillsdale, NJ: Erlbaum.

Lahav, R. 1993: What neuropsychology tells us about consciousness. *Philosophy of Science*, 60, 67–85.

Langton, C. G. 1995: *Artificial Life*. Cambridge, Mass.: MIT Press.

Lashley, K. S. 1929: *Brain Mechanisms and Intelligence*. Chicago: University of Chicago Press.

—— 1950: In search of the engram. *Symposia of the Society for Experimental Biology*, 4, 454–82.

—— 1951: The problem of serial order in behavior. In L. A. Jeffries (ed.), *Cerebral Mechanisms in Behavior: The Hixon Symposium*, New York: Wiley, 112–36.

Lee, D. N. 1976: A theory of the visual control of breaking based on information about time to collision. *Perception*, 5, 437–59.

Lenneberg, E. H. 1967: *Biological Foundations of Language*. New York: Wiley.

Lettvin, J. Y., Maturana, H. R., McCulloch, W. S. and Pitts, W. H. 1959: What the frog's eye tells the frog's brain. *Proceedings of the IRE*, 47, 1940–51.

Lindsay, P. and Norman, D. A. 1972: *Human Information Processing*. New York: Academic Press.

Lounsbury, F. G. 1956: A semantic analysis of the Pawnee kinship usage. *Language*, 32, 158–94.

Marr, D. 1969: A theory of the cerebral cortex. *Journal of Physiology*, 202, 437–70.

—— 1982: *Vision*. San Francisco: Freeman.

McCarthy, J. and Hayes, P. J. 1969: Some philosophical problems from the standpoint of artificial intelligence. In D. Michie (ed.), *Machine Intelligence*, vol. 4, Edinburgh: Edinburgh University Press, 463–503.

*McCorduck, P. 1979: *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence*. San Francisco: Freeman.

McCulloch, W. S. and Pitts, W. H. 1943: A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, 115–33.

Miller, G. A. 1956: The magic number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63, 81–97.

—— 1962: Some psychological studies of grammar. *American Psychologist*, 17, 748–62.

—— 1979: *A Very Personal History*. Cambridge, Mass.: MIT Center for Cognitive Science, Occasional Paper no. 1.

Miller, G. A. and Selfridge, J. A. 1950: Verbal context and the recall of meaningful material. *American Journal of Psychology*, 63, 176–85.

Miller, G. A., Galanter, E. and Pribram, K. 1960: *Plans and the Structure of Behavior*. New York: Holt.

Milner, B. 1965: Memory disturbance after bilateral hippocampal lesions. In P. M. Milner and S. E. Glickman (eds), *Cognitive Processes and the Brain*, Patterson, NJ: Van Nostrand, 97–111.

Minsky, M. (ed.) 1968: *Semantic Information Processing*. Cambridge, Mass.: MIT Press.

—— 1975: A framework for representing knowledge. In P. H. Winston (ed.), *The Psychology of Computer Vision*, New York: McGraw-Hill, 211–79.

Minsky, M. and Papert, S. 1969: *Perceptrons*. Cambridge, Mass.: MIT Press.

Mishkin, M., Ungerleider, L. G. and Macko, K. A. 1983: Object vision and spatial vision: two cortical pathways. *Trends in Neuroscience*, 6, 414–17.

Moore, O. K. and Anderson, S. B. 1954: Modern logic and tasks for experiments on problem solving behavior. *Journal of Psychology*, 38, 151–60.

Neisser, U. 1967: *Cognitive Psychology*. New York: Appleton-Century-Crofts.

—— 1976: *Cognition and Reality*. San Francisco: Freeman.

—— 1988: Cognitive recollections. In W. Hirst (ed.), *The Making of Cognitive Science: Essays in Honor of George A. Miller*, New York: Cambridge University Press, 81–8.

Newell, A. 1980: Physical symbol systems. *Cognitive Science*, 4, 135–83.

Newell, A. and Simon, H. 1956: The logic theory machine. *IRE Transactions on Information Theory*, 3, 61–79.

—— —— 1972: *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.

*Newmeyer, F. J. 1986: *Linguistic Theory in America*, 2nd edn. New York: Academic Press.

Norman, D. A. 1981: Preface to D. A. Norman (ed.), *Perspectives on Cognitive Science*, Norwood, NJ: Ablex.

Norman, D. A. and Levelt, W. J. 1988: Life at the Center. In W. Hirst (ed.), *The Making of Cognitive Science: Essays in Honor of George A. Miller*, New York: Cambridge University Press, 100–9.

Norman, D. A., Rumelhart, D. E. and the LNR Research Group 1975: *Explorations in Cognition*. San Francisco: Freeman.

Osgood, C. E. 1953: *Method and Theory in Experimental Psychology*. New York: Oxford University Press.

Osgood, C. E. and Sebeok, T. A. (eds) 1954: *Psycholinguistics: A Survey of Theory and Research Problems*, Indiana University Publications in Anthropology and Linguistics, Memoir 10 of *International Journal of American Linguistics*. Bloomington: Indiana University Press.

Osgood, C. E., Suci, G. J. and Tannenbaum, P. H. 1957: *The Measurement of Meaning*. Urbana: University of Illinois Press.

*Osherson, D. N. (ed.) 1995–1997: *An Invitation to Cognitive Science*, 4 vols. Cambridge, Mass.: MIT Press.

Penfield, W. and Rasmussen, T. 1950: *The Cerebral Cortex of Man: A Clinical Study of Localization of Function*. New York: Macmillan.

Peterson, L. R. and Peterson, M. J. 1959: Short-term retention of individual verbal items. *Journal of Experimental Psychology*, 58, 193–8.

Pitts, W. and McCulloch, W. S. 1947: How we know universals: the perception of auditory and visual forms. *Bulletin of Mathematical Biophysics*, 9, 127–47.

Port, R. and van Gelder, T. 1995: *Mind as Motion*. Cambridge, Mass.: MIT Press.

Posner, M. I. and McLeod, P. 1982: Information processing models: in search of elementary operations. *Annual Review of Psychology*, 33, 477–514.

Posner, M. I. and Raichle, M. E. 1994: *Images of Mind*. San Francisco: Freeman.

Posner, M. I., Pea, R. and Volpe, B. 1982: Cognitive-neuroscience: developments toward a science of synthesis. In J. Mehler, E. C. T. Walker, and M. Garrett (eds), *Perspectives on Mental Representation*, Hillsdale, NJ: Erlbaum, 251–76.

Posner, M. I., Petersen, S. E., Fox, P. T. and Raichle, M. E. 1988: Localization of cognitive operations in the human brain. *Science*, 240, 1627–31.

Prince, A. and Smolensky, P., in press: *Optimality Theory: Constraint Interaction in Generative Grammar*. Cambridge, Mass.: MIT Press.

Putnam, H. 1960: Minds and machines. In S. Hook (ed.), *Dimensions of Mind*, New York: New York University Press, 148–79.

—— 1967: Psychological predicates. In W. Capitan and D. Merrill (eds), *Art, Mind, and Religion*, Pittsburgh: University of Pittsburgh Press, 37–48.

Pylyshyn, Z. W. 1973: What the mind's eye tells the mind's brain: a critique of mental imagery. *Psychological Bulletin*, 80, 1–24.

—— 1981: The imagery debate: analogue media versus tacit knowledge. *Psychological Review*, 87, 16–45.

—— 1983: Information science: its roots and relations as viewed from the perspective of cognitive science. In F. Machlup and U. Mansfield (eds), *The Study of Information: Interdisciplinary Messages*, New York: John Wiley and Sons, 63–80.

Quillian, M. R. 1968: Semantic memory. In M. Minsky (ed.), *Semantic Information Processing*, Cambridge, Mass.: MIT Press, 227–70.

Ratliff, F. and Hartline, H. K. 1959: The response of limulus optic nerve fibers to patterns of illumination on the receptor mosaic. *Journal of General Physiology*, 42, 1241–55.

Rips, L. J., Shoben, E. J. and Smith, E. E. 1973: Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, 12, 1–20.

Rochester, N., Holland, J. H., Haibt, L. H. and Duda, W. L. 1956: Tests on a cell assembly theory of the action of the brain, using a large digital computer. *IRE Transactions on Information Theory*, 2, 80–93.

Rosenblatt, F. 1962: *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Washington, DC: Spartan Books.

Rosenblueth, A., Wiener, N. and Bigelow, J. 1943: Behavior, purpose, and teleology. *Philosophy of Science*, 10, 18–24.

Rumelhart, D. E., Hinton, G. E. and Williams, R. J. 1986: Learning representations by back-propagating errors. *Nature*, 323, 533–6.

Rumelhart, D. E., McClelland, J. L. and the PDP Research Group 1986: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, vol. 1: *Foundations*. Cambridge, Mass.: MIT Press.

Schank, R. C. and Abelson, R. P. 1977: *Scripts, Plans, Goals and Understanding*. Hillsdale, NJ: Erlbaum.

Schwartz, E. L. (ed.) 1990: *Computational Neuroscience*. Cambridge, Mass.: MIT Press.

Searle, J. 1980: Minds, brains, and programs. *Behavioral and Brain Sciences*, 3, 417–57.

Shannon, C. E. 1948: A mathematical theory of communication. *Bell System Technical Journal*, 27, 379–423, 623–56.

—— 1950: A chess-playing machine. *Scientific American*, 182(2), 48–51.

—— 1951: Prediction and entropy of printed English. *Bell System Technical Journal*, 30, 50–64.

Shepard, R. N. and Metzler, J. 1971: Mental rotation of three-dimensional objects. *Science*, 171, 701–3.

Shortliffe, E. H. 1976: *Mycin: Computer Based Medical Consultations*. New York: Elsevier.

Skinner, B. F. 1938: *The Behavior of Organisms: An Experimental Analysis*. New York: Appleton-Century.

—— 1957: *Verbal Behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Smolensky, P. 1988: On the proper treatment of connectionism. *Behavior and Brain Sciences*, 11, 1–74.

Sperling, G. 1960: *The Information Available in Brief Visual Presentations*, Psychological Monographs, 74 (whole no. 498).

Sternberg, S. 1969: The discovery of processing stages: extension of Donders' method. *Acta Psychologica*, 30, 276–315.

Turing, A. M. 1936: On computable numbers, with an application to the Entscheidungs-Problem. *Proceedings of the London Mathematical Society*, 2nd ser., 42, 230–65.

—— 1950: Computing machinery and intelligence. *Mind*, 59, 433–60.

Ullman, S. 1980: Against direct perception. *Behavioral and Brain Sciences*, 3, 373–416.

Valenstein, E. S. 1973: *Brain Control: A Critical Examination of Brain Stimulation and Psychosurgery*. New York: Wiley Interscience.

van Essen, D. C. and Anderson, C. H. 1990: Information processing strategies and pathways in the primate retina and visual cortex. In S. F. Zornetzer, J. L. Davis, and C. Lau (eds), *An Introduction to Neural and Electronic Networks*, New York: Academic Press, 43–72.

van Gelder, T. 1995: What might cognition be, if not computation? *Journal of Philosophy*, 92, 345–81.

Watson, J. 1913: Psychology as a behaviorist views it. *Psychological Review*, 20, 158–77.

Waugh, N. C. and Norman, D. A. 1965: Primary memory. *Psychological Review*, 72, 89–104.

Weizenbaum, J. 1976: *Computer Power and Human Reason: From Judgment to Calculation*. San Francisco: Freeman.

Wernicke, C. 1874: *Der aphasiche Symptomcomplex: Eine psychologische Studie auf anatomischer Basis*. Breslau: Cohen and Weigert.

Wertheimer, M. 1912: Experimentelle Studien über das Sehen von Bewegungen. *Zeitschrift für Psychologie*, 61, 161–265.

Whybrow, P. 1996: *A Mood Apart: Depression, Manic Depression, and Other Afflictions of the Self*. New York: Basic Books.

Wiener, N. 1948: *Cybernetics: Or, Control and Communication in the Animal Machine*. New York: Wiley.

Willner, P. (ed.) 1991: *Behavioral Models in Psychopharmacology: Theoretical, Industrial, and Clinical Perspectives*. Cambridge: Cambridge University Press.

Winograd, T. 1972: *Understanding Natural Languages*. New York: Academic Press.

Zeki, S. M. 1976: The functional organization of projections from striate to prestriate visual cortex in rhesus monkey. *Cold Spring Harbor Symposium on Quantitative Biology*, 40, 591–600.