

Other Books by Howard Gardner

The Quest for Mind (1973; second edition, 1981)

The Arts and Human Development (1973)

The Shattered Mind (1975)

Developmental Psychology (1978; second edition, 1982)

Artful Scribbles: The Significance of Children's Drawings (1980)

Art, Mind, and Brain: A Cognitive Approach to Creativity (1982)

Frames of Mind: The Theory of Multiple Intelligences (1983)

THE MIND'S NEW SCIENCE

A History of the
Cognitive Revolution

HOWARD GARDNER

With a New Epilogue by the Author:
Cognitive Science After 1984

Basic Books, Inc., Publishers / New York

I

Introduction: What the *Meno* Wrought

One thing I would fight for to the end, both in word and deed if I were able—that if we believed that we must try to find out what is not known, we should be better and braver and less idle than if we believed that what we do not know it is impossible to find out and that we need not even try.

—SOCRATES, *The Meno*

The safest general characterization of the European philosophical tradition is that it consists in a series of footnotes to Plato.

—ALFRED NORTH WHITEHEAD

The Greek Agenda

In the *Meno*, a Platonic dialogue, Socrates persistently questions a young slave about his knowledge of geometry. At first the slave appears quite knowledgeable, readily asserting that a square composed of sides two feet in length contains four square feet. But when, in response to a problem posed by Socrates, the slave indicates that a figure of eight square feet contains sides four feet long, Socrates demonstrates that the boy is thoroughly confused and does not realize that the length of the side must be the square root of eight.

The centerpiece of the dialogue features many questions and responses in the approved Socratic manner. Through this interchange, the philosopher ultimately succeeds in drawing out from the boy the knowledge that a square with a four-foot side would actually be sixteen square feet—that is, twice as great an area than he had supposed; and the knowledge that one can, by geometric maneuvers, inscribe a square that is actually eight square feet within this larger square. In so doing, Socrates has demonstrated to his satisfaction, and to the satisfaction of the slave's master, Menon, that the youth possesses within him all of the knowledge necessary to compute the various geometrical relationships in question.

At issue in this Platonic dialogue was far more than an exploration of the extent of knowledge possessed by a single slave boy. Here, for perhaps the first time in human intellectual history, was an extended rumination on the nature of knowledge: where does it come from, what does it consist of, how is it represented in the human mind? And, for good measure, there was also proposed a specific—if ultimately highly controversial—theory of human knowledge.

According to Plato (and, presumably, Socrates as well), the domain of knowledge par excellence inhered in mathematics and the exact sciences it had spawned. Indeed, the purest forms of knowledge were idealized forms or archetypes which can merely be glimpsed in mundane reality. An understanding of geometrical matters—indeed, of all matters of genuine knowledge—was already implanted in the human soul at birth. The task in instruction, as demonstrated in the dialogue of the *Meno*, was simply to bring this innate knowledge to conscious awareness.

The Greeks' interest in the nature of knowledge, no less than their particular contentious theories and evocative images, continued to reverberate through the Western intellectual tradition. Aristotle's version was the principal cornerstone of the Middle Ages, when discussions about knowledge were principally the purview of theologians. Then, during the Renaissance and Enlightenment periods, philosophers continued the discussions and began to draw regularly on findings obtained in the newly emerging empirical sciences. Such thinkers as Descartes, Locke, and Kant dealt comfortably with theoretical and empirical issues concerning knowledge, and the Neapolitan scholar Giambattista Vico even christened a New Science (*Scienza Nova*) to deal with these and related matters. By the end of the nineteenth century, there had been a proliferation of new sciences and philosophical specialties, several of which purported to deal with the nature of the human mind.

Today, armed with tools and concepts unimaginable even a century ago, a new cadre of thinkers called cognitive scientists has been investigating many of the same issues that first possessed the Greeks some twenty-

Introduction: What the Meno Wrought

five hundred years ago. Like their earlier counterparts, cognitive scientists today ask what it means to know something and to have accurate beliefs, or to be ignorant or mistaken. They seek to understand what is known—the objects and subjects in the external world—and the person who knows—his* perceptual apparatus, mechanisms of learning, memory, and rationality. They ponder the sources of knowledge: where does it come from, how is it stored and tapped, how might it be lost? They are curious about the differences among individuals: who learns early or with difficulty; what can be known by the child, the inhabitant of a preliterate society, an individual who has suffered brain damage, or a mature scientist?

Further, cognitive scientists, again as did the Greeks, conjecture about the various vehicles of knowledge: what is a form, an image, a concept, a word; and how do these "modes of representation" relate to one another? They wonder about the priorities of specific sense organs as against a central "general understanding" or "common sense." They reflect on language, noting the power and traps entailed in the use of words and their possible predominant influence over thoughts and beliefs. And they speculate at length on the nature of the very activity of knowing: why do we want to know, what are the constraints on knowing, and what are the limits of scientific knowledge about human knowing?

This "new science," thus, reaches back to the Greeks in the commitment of its members to unraveling the nature of human knowledge. At the same time, however, it is radically new. Proceeding well beyond armchair speculation, cognitive scientists are fully wedded to the use of empirical methods for testing their theories and their hypotheses, of making them susceptible to disconfirmation. Their guiding questions are not just a rehash of the Greek agenda: new disciplines, like artificial intelligence, have arisen; and new questions, like the potential of man-made devices to think, stimulate research. Moreover, cognitive scientists embrace the most recent scientific and technological breakthroughs in a variety of disciplines. Most central to their undertaking is the computer—that creation of the mid-twentieth century that holds promise for changing our conceptions of the world in which we live and our picture of the human mind.

Definition and Scope of Cognitive Science

In the course of proposing and founding a new field of knowledge, many individuals will formulate their own definitions. Indeed, since the term *cognitive science* first began to be bandied about in the early 1970s, dozens of scientists have attempted to define the nature and scope of the

*For ease of exposition, the pronoun *he* is used in its generic sense throughout this book.

field (see, for example, Bruner 1983; Collins 1977; Mandler 1981; Miller 1979; Norman 1980; Rumelhart 1982). It therefore becomes important for me at the outset to state what I take cognitive science to be.

I define cognitive science as a contemporary, empirically based effort to answer long-standing epistemological questions—particularly those concerned with the nature of knowledge, its components, its sources, its development, and its deployment. Though the term *cognitive science* is sometimes extended to include all forms of knowledge—animate as well as inanimate, human as well as nonhuman—I apply the term chiefly to efforts to explain human knowledge. I am interested in whether questions that intrigued our philosophical ancestors can be decisively answered, instructively reformulated, or permanently scuttled. Today cognitive science holds the key to whether they can be.

Of the various features or aspects generally associated with cognitive-scientific efforts, I consider five to be of paramount importance. Not every cognitive scientist embraces every feature, of course; but these features can be considered symptomatic of the cognitive-scientific enterprise. When all or most are present, one can assume that one is dealing with cognitive science; when few, if any, are present, one has fallen outside my definition of cognitive science. These features will be introduced more formally at the end of chapter 3 and will be revisited repeatedly throughout the book, but it is important to make an initial acquaintance with them at this point.

First of all, there is the belief that, in talking about human cognitive activities, it is necessary to speak about mental representations and to posit a level of analysis wholly separate from the biological or neurological, on the one hand, and the sociological or cultural, on the other.

Second, there is the faith that central to any understanding of the human mind is the electronic computer. Not only are computers indispensable for carrying out studies of various sorts, but, more crucially, the computer also serves as the most viable model of how the human mind functions.

While the first two features incorporate the central beliefs of current cognitive science, the latter three concern methodological or strategic characteristics. The third feature of cognitive science is the deliberate decision to de-emphasize certain factors which may be important for cognitive functioning but whose inclusion at this point would unnecessarily complicate the cognitive-scientific enterprise. These factors include the influence of affective factors or emotions, the contribution of historical and cultural factors, and the role of the background context in which particular actions or thoughts occur.

As a fourth feature, cognitive scientists harbor the faith that much is to be gained from interdisciplinary studies. At present most cognitive

Introduction: What the Meno Wrought

scientists are drawn from the ranks of specific disciplines—in particular, philosophy, psychology, artificial intelligence, linguistics, anthropology, and neuroscience (I shall refer to these disciplines severally as the “cognitive sciences”). The hope is that some day the boundaries between these disciplines may become attenuated or perhaps disappear altogether, yielding a single, unified cognitive science.

A fifth and somewhat more controversial feature is the claim that a key ingredient in contemporary cognitive science is the agenda of issues, and the set of concerns, which have long exercised epistemologists in the Western philosophical tradition. To my mind, it is virtually unthinkable that cognitive science would exist, let alone assume its current form, had there not been a philosophical tradition dating back to the time of the Greeks.

Purpose and Plan of This Book

I have chosen to write a book on cognitive science because I consider this area to be the most exciting new line of inquiry undertaken by scientists in the past few decades. Whether it will ultimately achieve all of its objectives, no one can say at this point; but this seems an opportune time to present a history and a current assessment. For contemporaries present during the opening decades of cognitive science, I hope to convey something of the enthusiasm I have noted, the difficulties that are being confronted, and the nature of the research enterprises in which investigators are presently engaged.

My history has two components. The first consists of the various interdisciplinary conversations and projects that took place in this century—both those preceding and those surrounding the unofficial launching of cognitive science in the mid-1950s. I relate the founding of cognitive science in the next two chapters of the book. The second component—spanning chapters 4 through 9—consists of brief targeted histories of each of the six aforementioned fields of cognitive science. (Other disciplines, such as sociology or economics, might have been added; the “borderline” disciplines of anthropology and neuroscience might have been eliminated; but I believe that the major points about cognitive science are made effectively by these six fields.) In my view, a brief targeted history of each of the several cognitive sciences serves as an optimal introduction to the principal issues of today, to the ways in which they are currently approached and explored, and to the lines of work likely to be undertaken in the future.

I have built each historical chapter around one or two major themes, which have been selected to convey a feeling for the kinds of issues that

have recurred and the kinds of approaches that are especially central within a particular field. For example, in philosophy I trace the perennial dispute between those of a rationalist persuasion (who view the mind as actively organizing experiences on the basis of pre-existing schemes); and those of an empiricist bent (who treat mental processes as a reflection of information obtained from the environment). In anthropology I survey various attempts over the years to compare the thought of primitive peoples with that exhibited by typical individuals in modern Western society. Approaching these same fields from a methodological point of view, I raise the questions whether philosophy will eventually come to be supplanted by an empirically based cognitive science, and whether anthropology can (or even should) ever transcend the individual case study.

Of course, such organizing themes can only scratch the surface of the complex territory that underlies any scientific discipline. Still I hope that through such themes I can convey how a linguist views an issue, what a psychologist deems a problem (and a solution), which conceptions of process obtain in neuroscience and artificial intelligence. Only through such an immersion in the daily (and yearly) concerns of a cognitive scientist drawn from a particular discipline can one appreciate the possibilities—and the difficulties—that arise when workers from different fields collaborate in cognitive-scientific research. In the end I will in each case take stock and indicate where things stand with reference to the principal lines of contention in a particular cognitive science—a discussion that will, in turn, suggest some of the principal factors that have stimulated cognitive scientists to join forces.

While each of the histories stands on its own, their juxtaposition points up fascinating and difficult-to-anticipate parallels. Scientific fields hardly develop in a vacuum: such disparate factors as the dissemination of Darwin's pivotal writings, the outbreak of wars, the rise of great universities have had reverberations—and sometimes cataclysmic ones—across apparently remote fields, which may well have had little direct contact with one another. For the most part, I shall simply allow these parallels to emerge, but at the beginning of part III I shall specify certain historical forces that seem to have exerted influence across a range of cognitive sciences.

Having taken the measure of the individual cognitive sciences, I turn in the third part of the book to review ongoing work that is quintessentially cognitive-scientific. Thus, in chapters 10 to 13, the focus shifts from work within a traditional discipline to those lines of research that stand most squarely at the intersection of a number of disciplines and therefore can be considered prototypical of a single, unified cognitive science. I have sought to identify work that is of the highest quality: if cognitive science

is to be assessed as an intellectual enterprise, it ought to be judged by the most outstanding instances.

There is a common structure to these four essays on current cognitive-scientific work. Consistent with my claim that cognitive science seeks to elucidate basic philosophical questions, each chapter begins with a perennial epistemological issue. For example, in chapter 10, I describe work on how we perceive the world; in chapter 13, I review competing claims on the extent of human rationality. Across chapters 10 to 13, there is a progression from those issues that seem most circumscribed to those that are most global. Not surprisingly, the most confident answers exist for the delimited questions, while the global topics remain ringed by unresolved questions.

My personal reflections on cognitive science are reserved for the final chapter. There I revisit the major themes of cognitive science in light of the histories sketched and the interdisciplinary work reviewed. I also discuss two themes that emerge from the inquiry and that will be introduced at greater length in chapter 3: the computational paradox and the cognitive challenge. In my view, the future of cognitive science rests on how the computational paradox is resolved and on how the cognitive challenge is met.

One might say that cognitive science has a very long past but a relatively short history. The reason is that its roots go back to classical times, but it has emerged as a recognized pursuit only in the last few decades. Indeed, it seems fair to maintain that the various components that gave rise to cognitive science were all present in the early part of the century, and the actual birthdate occurred shortly after mid-century. Just why cognitive science arose when it did in the form it did will constitute my story in the remainder of part I.

Cognitive Science: The First Decades

A Consensual Birthdate

Seldom have amateur historians achieved such consensus. There has been nearly unanimous agreement among the surviving principals that cognitive science was officially recognized around 1956. The psychologist George A. Miller (1979) has even fixed the date, 11 September 1956.

Why this date? Miller focuses on the Symposium on Information Theory held at the Massachusetts Institute of Technology on 10–12 September 1956 and attended by many leading figures in the communication and the human sciences. The second day stands out in Miller's mind because of two featured papers. The first, presented by Allen Newell and Herbert Simon, described the "Logic Theory Machine," the first complete proof of a theorem ever carried out on a computing machine. The second paper, by the young linguist Noam Chomsky, outlined "Three Models of Language." Chomsky showed that a model of language production derived from Claude Shannon's information-theoretical approach could not possibly be applied successfully to "natural language," and went on to exhibit his own approach to grammar, based on linguistic transformations. As Miller recalls, "Other linguists had said language has all the formal precisions of mathematics, but Chomsky was the first linguist to make good on the claim. I think that was what excited all of us" (1979, p. 8). Not incidentally, that day George Miller also delivered a seminal paper, outlining his claim that the capacity of human short-term memory is limited to approximately seven entries. Miller summed up his reactions:

I went away from the Symposium with a strong conviction, more intuitive than rational, that human experimental psychology, theoretical linguistics, and computer simulation of cognitive processes were all pieces of a larger whole, and that the future would see progressive elaboration and coordination of their shared concerns. . . . I have been working toward a cognitive science for about twenty years beginning before I knew what to call it. (1979, p. 9)

Miller's testimony is corroborated by other witnesses. From the ranks of psychology, Jerome Bruner declares, "New metaphors were coming into being in those mid-1950s and one of the most compelling was that of computing. . . . My "Generation" created and nurtured the Cognitive Revolution—a revolution whose limits we still cannot fathom" (1983, pp. 274, 277). Michael Posner concludes, "This mix of ideas about cognition was ignited by the information processing language that arrived in psychology in the early 1950s" (Posner and Shulman 1979, p. 374). And George Mandler suggests:

For reasons that are obscure at present, the various tensions and inadequacies of the first half of the twentieth century cooperated to produce a new movement in psychology that first adopted the label of information processing and after became known as modern cognitive psychology. And it all happened in the five year period between 1955 and 1960. Cognitive science started during that five year period, a happening that is just beginning to become obvious to its practitioners. (1981, p. 9)

Finally, in their history of the period, computer scientists Allen Newell and Herbert Simon declare:

Within the last dozen years a general change in scientific outlook has occurred, consonant with the point of view represented here. One can date the change roughly from 1956: in psychology, by the appearance of Bruner, Goodnow, and Austin's *Study of Thinking* and George Miller's "The magical number seven"; in linguistics, by Noam Chomsky's "Three models of language"; and in computer science, by our own paper on the Logical Theory Machine. (1972, p. 4)

This impressive congruence stresses a few seminal publications, emanating (not surprisingly perhaps) from the same small group of investigators. In fact, however, the list of relevant publications is almost endless. As far as general cognitive scientific publications are concerned, John von Neumann's posthumous book, *The Computer and the Brain* (1958), should head the list. In this book—actually a set of commissioned lectures which von Neumann became too ill to deliver—the pioneering computer scientist developed many of the themes originally touched upon in his Hixon Symposium contribution. He included a discussion of various kinds of comput-

ers and analyzed the idea of a program, the operation of memory in computers, and the possibility of machines that replicate themselves.

Relevant research emanated from each of the fields that I have designated as contributing cognitive sciences.* The witnesses I have just quoted noted the principal texts in the fields of psychology, linguistics, and artificial intelligence, and many more entries could be added. Neuroscientists were beginning to record impulses from single neurons in the nervous system. At M.I.T., Warren McCulloch's research team, led by the neurophysiologists Jerome Lettvin and Humberto Maturana, recorded from the retina of the frog. They were able to show that neurons were responsive to extremely specific forms of information such as "bug-like" small dark spots which moved across their receptive fields, three to five degrees in extent. Also in the late 1950s, a rival team of investigators, David Hubel and Torsten Wiesel at Harvard, began to record from cells in the visual cortex of the cat. They located nerve cells that responded to specific information, including brightness, contrast, binocularity, and the orientation of lines. These lines of research, eventually honored in 1981 by a Nobel Prize, called attention to the extreme specificity encoded in the nervous system.

The mid 1950s were also special in the field of anthropology. At this time, the first publications by Harold Conklin, Ward Goodenough, and Floyd Lounsbury appeared in the newly emerging field of cognitive anthropology, or ethnosemantics. Researchers undertook systematic collection of data concerning the naming, classifying, and concept-forming abilities of people living in remote cultures, and then sought to describe in formal terms the nature of these linguistic and cognitive practices. These studies documented the great variety of cognitive practices found around the world, even as they strongly suggested that the relevant cognitive processes are similar everywhere.

In addition, in the summer of 1956, a group of young scholars, trained in mathematics and logic and interested in the problem-solving potentials of computers, gathered at Dartmouth College to discuss their mutual interests. Present at Dartmouth were most of the scholars working in what came to be termed "artificial intelligence," including the four men generally deemed to be its founding fathers: John McCarthy, Marvin Minsky, Allen Newell, and Herbert Simon. During the summer institute, these scientists, along with other leading investigators, reviewed ideas for programs that would solve problems, recognize patterns, play games, and reason logically, and laid out the principal issues to be discussed in coming years. Though no synthesis emerged from these discussions, the participants seem to have set up a permanent kind of "in group" centered at the

* Full bibliographical references to these lines of research will be provided at appropriate points in the text.

M.I.T., Stanford, and Carnegie-Mellon campuses. To artificial intelligence, this session in the summer of 1956 was as central as the meeting at M.I.T. among communication scientists a few months later.

Scholars removed from empirical science were also pondering the implications of the new machines. Working at Princeton, the American philosopher Hilary Putnam (1960) put forth an innovative set of notions. As he described it, the development of Turing-machine notions and the invention of the computer helped to solve—or to dissolve—the classical mind-body problem. It was apparent that different programs, on the same or on different computers, could carry out structurally identical problem-solving operations. Thus, the logical operations themselves (the "software") could be described quite apart from the particular "hardware" on which they happened to be implemented. Put more technically, the "logical description" of a Turing machine includes no specification of its physical embodiment.

The analogy to the human system and to human thought processes was clear. The human brain (or "bodily states") corresponded to the computational hardware; patterns of thinking or problem solving ("mental states") could be described entirely separately from the particular constitution of the human nervous system. Moreover, human beings, no less than computers, harbored programs; and the same symbolic language could be invoked to describe programs in both entities. Such notions not only clarified the epistemological implications of the various demonstrations in artificial intelligence; they also brought contemporary philosophy and empirical work in the cognitive sciences into much closer contact.

One other significant line of work, falling outside cognitive science as usually defined, is the ethological approach to animal behavior which had evolved in Europe during the 1930s and 1940s thanks to the efforts of Konrad Lorenz (1935) and Niko Tinbergen (1951). At the very time that American comparative psychologists were adhering closely to controlled laboratory settings, European ethologists had concluded that animals should be studied in their natural habitat. Observing carefully under these naturalistic conditions, and gradually performing informal experiments on the spot, the ethologists revealed the extraordinary fit between animals and their natural environment, the characteristic *Umwelt* (or world view) of each species, and the particular stimuli (or releasers) that catalyze dramatic developmental milestones during "critical" or "sensitive" periods. Ethology has remained to some extent a European rather than an American specialty. Still, the willingness to sample wider swaths of behavior in naturally occurring settings had a liberating influence on the types of concept and the modes of exploration that came to be tolerated in cognitive studies.

The 1960s: Picking Up Steam

The seeds planted in the 1950s sprouted swiftly in the next decade. Governmental and private sources provided significant financial support. Setting the intellectual tone were the leading researchers who had launched the key lines of study of the 1950s, as well as a set of gifted students who were drawn to the cognitive fields, much in the way that physics and biology had lured the keenest minds of earlier generations. Two principal figures in this "selling of cognition" were Jerome Bruner and George Miller, who in 1960 founded at Harvard the Center for Cognitive Studies. The Center, as story has it, began when these two psychologists approached the dean of the faculty, McGeorge Bundy, and asked him to help create a research center devoted to the nature of knowledge. Bundy reportedly responded, "And how does that differ from what Harvard University does?" (quoted in Bruner 1983, p. 123). Bundy gave his approval, and Bruner and Miller succeeded in getting funds from the Carnegie Corporation, whose president at that time, the psychologist John Gardner, was sympathetic to new initiatives in the behavioral sciences.

Thereafter, for over ten years, the Harvard Center served as a locale where visiting scholars were invited for a sabbatical, and where graduate and postdoctorate students flocked in order to sample the newest thinking in the cognitive areas. A list of visitors to the Center reads like a Who's Who in Cognitive Science: nearly everyone visited at one time or another, and many spent a semester or a year in residence. And while the actual projects and products of the Center were probably not indispensable for the life of the field, there is hardly a younger person in the field who was not influenced by the Center's presence, by the ideas that were bandied about there, and by the way in which they were implemented in subsequent research. Indeed, psychologists Michael Posner and Gordon Shulman (1979) locate the inception of the cognitive sciences at the Harvard Center.

During the 1960s, books and other publications made available the ideas from the Center and from other research sites. George Miller—together with his colleagues Karl Pribram, a neuroscientist, and Eugene Galanter, a mathematically oriented psychologist—opened the decade with a book that had a tremendous impact on psychology and allied fields—a slim volume entitled *Plans and the Structure of Behavior* (1960). In it the authors sounded the death knell for standard behaviorism with its discredited reflex arc and, instead, called for a cybernetic approach to behavior in terms of actions, feedback loops, and readjustments of action in the light

Cognitive Science: The First Decades

of feedback. To replace the reflex arc, they proposed a unit of activity called a "TOTE unit" (for "Test-Operate-Test-Exit"): an important property of a TOTE unit was that it could itself be embedded within the hierarchical structure of an encompassing TOTE unit. As a vehicle for conceptualizing such TOTE units, the authors selected the computer with its programs. If a computer could have a goal (or a set of goals), a means for carrying out the goal, a means for verifying that the goal has been carried out, and then the option of either progressing to a new goal or terminating behavior, models of human beings deserved no less. The computer made it legitimate in theory to describe human beings in terms of plans (hierarchically organized processes), images (the total available knowledge of the world), goals, and other mentalistic conceptions; and by their ringing endorsement, these three leading scientists now made it legitimate in practice to abandon constricted talk of stimulus and response in favor of more open-ended, interactive, and purposeful models.

The impact of this way of thinking became evident a few years later when textbooks in cognitive psychology began to appear. By far the most influential was *Cognitive Psychology* by the computer-literate experimental psychologist Ulric Neisser (1967). Neisser put forth a highly "constructive" view of human activity. On his account, all cognition, from the first moment of perception onward, involves inventive analytic and synthesizing processes. He paid tribute to computer scientists for countenancing talk of an "executive" and to information scientists for discussing accession, processing, and transformation of data. But at the same time, Neisser resisted uncritical acceptance of the computer-information form of analysis. In his view, objective calculation of how many bits of information can be processed is not relevant to psychology, because human beings are selective in their attention as a pure channel such as a telephone cannot be. Neisser expressed similar skeptical reservations about the claims surrounding computer programs:

None of [these programs] does even remote justice to the complexity of human mental processes. Unlike men, "artificially intelligent" programs tend to be single minded, undistractable, and unemotional. . . . This book can be construed as an extensive argument against models of this kind, and also against other simplistic theories of the cognitive processes. (1967, p. 9)

After Neisser, it was possible to buy the cognitive science approach in general and still join into vigorous controversies with "true believers."

Enthusiasts of the power of simulation were scarcely silent during this period. In his 1969 Compton lectures, *The Sciences of the Artificial*, Herbert Simon provided a philosophical exposition of his approach: as he phrased

it, both the computer and the human mind should be thought of as "symbol systems"—physical entities that process, transform, elaborate, and, in other ways, manipulate symbols of various sorts. And, in 1972, Allen Newell and Herbert Simon published their magnum opus, the monumental *Human Problem Solving*, in which they described the "general problem solver" programs, provided an explanation of their approach to cognitive studies, and included a historical addendum detailing their claims to primacy in this area of study.

Textbooks and books of readings were appearing in other subfields of cognitive science as well. An extremely influential collection was Jerry Fodor and Jerrold Katz's collection, *The Structure of Language* (1964), which anthologized articles representing the Chomskian point of view in philosophy, psychology, and linguistics, and attempted to document why this approach, rather than earlier forays into language, was likely to be the appropriate scientific stance. In artificial intelligence, Edward Feigenbaum and Julian Feldman put out a collection called *Computers and Thought* (1963), which presented many of the best-running programs of the era; while their collection had a definite "Carnegie slant," a rival anthology, *Semantic Information Processing*, edited by Marvin Minsky in 1968, emphasized the M.I.T. position. And, in the area of cognitive anthropology, in addition to influential writings by Kimball Romney and Roy D'Andrade (1964), Stephen Tyler's textbook *Cognitive Anthropology* made its debut in 1969.

But by 1969, the number of slots in short-term memory had been exceeded—without the benefit of chunking, one could no longer enumerate the important monographs, papers, and personalities in the cognitive sciences. (In fact, though my list of citations may seem distressingly long, I have really only scratched the surface of cognitive science, circa 1970.) There was tremendous activity in several fields, and a feeling of definite progress as well. As one enthusiastic participant at a conference declared:

We may be at the start of a major intellectual adventure: somewhere comparable to the position in which physics stood toward the end of the Renaissance, with lots of discoveries waiting to be made and the beginning of an inkling of an idea of how to go about making them. It turned out, in the case of the early development of modern physics that the advancement of the science involved developing new kinds of intellectual sophistication: new mathematics, a new ontology, and a new view of scientific method. My guess is that the same sort of evolution is required in the present case (and, by the way, in much the same time scale). Probably now, as then, it will be an uphill battle against obsolescent intellectual and institutional habits. (Sloan Foundation 1976, p. 10)

When the amount of activity in a field has risen to this point, with an aura of excitement about impending breakthroughs, human beings

often found some sort of an organization or otherwise mark the new enterprise. Such was happening in cognitive science in the early and middle 1970s. The moment was ripe for the coalescing of individuals, interests, and disciplines into an organizational structure.

The Sloan Initiative

At this time, fate intervened in the guise of a large New York-based private foundation interested in science—the Alfred P. Sloan Foundation. The Sloan Foundation funds what it terms "Particular Programs," in which it invests a sizable amount of money in an area over a few years' time, in the hope of stimulating significant progress. In the early 1970s, a Particular Program had been launched in the neurosciences: a collection of disciplines that explore the nervous system—ranging from neuropsychology and neurophysiology to neuroanatomy and neurochemistry. Researchers drawn from disparate fields were stimulated by such funding to explore common concepts and common organizational frameworks. Now Sloan was casting about for an analogous field, preferably in the sciences, in which to invest a comparable sum.

From conversations with officers of the Sloan Foundation, and from the published record, it is possible to reconstruct the principal events that led to the Sloan Foundation's involvement with cognitive science. In early 1975, the foundation was contemplating the support of programs in several fields; but by late 1975, a Particular Program in the cognitive sciences was the major one under active consideration. During the following year, meetings were held where major cognitive scientists shared their views. Possibly sensing the imminent infusion of money into the field, nearly every scientist invited by the Sloan Foundation managed to juggle his or her schedule to attend the meetings. Though there was certainly criticism voiced of the new cognitive science movement, most participants (who were admittedly interested parties) stressed the promise of the field and the need for flexible research and training support.

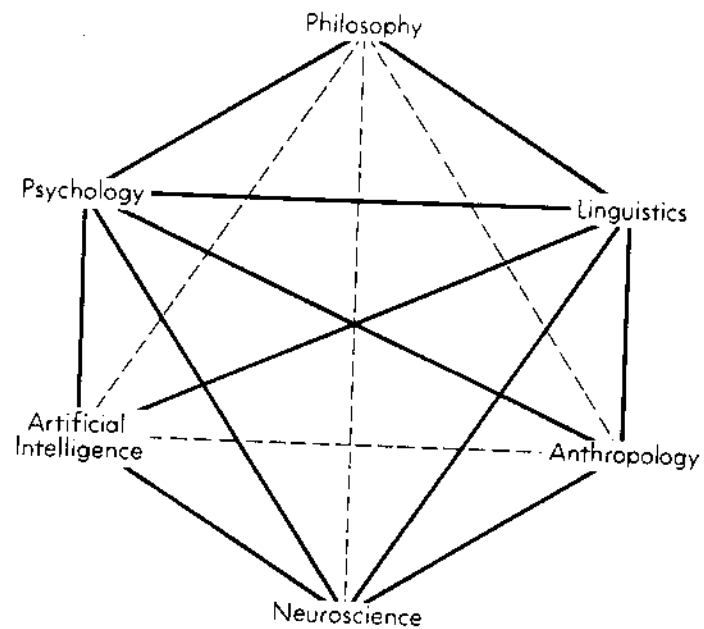
While recognizing that cognitive science was not as mature as neuroscience at the time of the foundation's commitment to the latter field, officers concluded that "nonetheless, there is every indication, confirmed by the many authorities involved in primary explorations, that many areas of the cognitive sciences are converging, and, moreover, there is a correspondingly important need to develop lines of communication

from area to area so that research tools and techniques can be shared in building a body of theoretical knowledge" (Sloan Foundation 1976, p. 6). After deliberating, the foundation decided to embark on a five-to-seven-year program, involving commitments of up to fifteen million dollars. (This commitment was ultimately increased to twenty million dollars.) The investment took the form, initially, of small grants to many research institutions and, ultimately, of a few large-scale grants to major universities.

Like the spur provided by the Macy Foundation a generation earlier, the Sloan Foundation's initiative had a catalytic effect on the field. As more than one person quipped, "Suddenly I woke up and discovered that I had been a cognitive scientist all of my life." In short order the journal *Cognitive Science* was founded—its first issue appearing in January 1977; and soon thereafter, in 1979, a society of the same name was founded. Donald Norman of the University of California in San Diego was instrumental in both endeavors. The society held its first annual meeting, amid great fanfare, in La Jolla, California, in August 1979. Programs, courses, newsletters, and allied scholarly paraphernalia arose around the country and abroad. There were even books about the cognitive sciences, including a popular account, *The Universe Within*, by Morton Hunt (1982) and my own historical essay, also supported by the Sloan Foundation.

Declaring the birth of a field had a bracing effect on those who discovered that they were in it, either centrally or peripherally, but by no means ensured any consensus, let alone appreciable scientific progress. Patrons are almost always necessary, though they do not necessarily suffice, to found a field or create a consensus. Indeed, tensions about what the field is, who understands it, who threatens it, and in what direction it ought to go were encountered at every phase of the Sloan Foundation's involvement (and have continued to be to this day).

Symptomatic of the controversy engendered by the Sloan Foundation's support of research in cognitive science was the reaction to a report commissioned by the foundation in 1978. This State of the Art Report (soon dubbed "SOAP" for short) was drafted by a dozen leading scholars in the field, with input from another score of advisers. In the view of the authors, "What has brought the field into existence is a common research objective: to discover the representational and computational capacities of the mind and their structural and functional representation in the brain" (1978, p. 6). The authors prepared a sketch of the interrelations among the six constituent fields—the cognitive hexagon, as it was labeled. Through the use of unbroken and broken lines, an effort was made to indicate the connections between fields which had already been forged, and to suggest the kinds of connection which could be but had not yet been effected.



Connections among the Cognitive Sciences

KEY: Unbroken lines = strong interdisciplinary ties
Broken lines = weak interdisciplinary ties

In my view, the authors of the SOAP document made a serious effort to survey principal lines of research and to provide a general charter for work in cognitive science, setting forth its principal assumptions. Then, using the example of how individuals from different cultures give names to colors, these authors illustrated how different disciplines combine their insights. (I'll flesh out this example of color naming in chapter 12.) However, the community-at-large adopted a distinctly negative view of the report. In fact, such virulent opposition was expressed by so many readers that, counter to original plans, the document was never published. I think this negative reaction came from the fact that each reader approached the document from the perspective of his or her own discipline and research program. In an effort to be reasonably ecumenical, the authors simply ensured that most readers would find their own work slighted. Moreover, there is as yet no agreed-upon research paradigm—no consensual set of assumptions or methods—and so cognitive scientists tend to project their own favorite paradigms onto the field as a whole. In view of these factors, it was probably not possible in 1978 to write a document that would have won the support of a majority of cognitive scientists.

It would be desirable, of course, for a consensus mysteriously to

emerge, thanks to the largesse of the Sloan Foundation, or for some latter-day Newton or Darwin to bring order into the field of cognitive science. In the absence, however, of either of these miraculous events, it is left to those of us who wish to understand cognitive science to come up with our own tentative formulation of the field. In the opening chapter of this book, I presented a working definition of cognitive science and alluded to five key components of the field. Now that I have sketched out some of the intellectual forces that led to the launching of cognitive science some three decades ago, I want to revisit these themes in somewhat more detail, in order to consider some of their implications as well as some of their problematic aspects. I will then conclude this introductory part by describing the paradox and the challenge standing at the center of contemporary cognitive science.

Key Features of Cognitive Science

In my own work I have found it useful to distinguish five features or "symptoms" of cognitive science: the first two of these represent the "core assumptions" of the field, while the latter three represent methodological or strategic features. Not only are these ideas common to most "strong versions" of cognitive science, but they also serve as specific points of contention for its critics. I shall list each of these characteristics and then indicate certain lines of criticism put forth by those most antagonistic to cognitive science. These criticisms (as voiced by their most vocal adherents) will be expanded upon at appropriate points in the book and reviewed in my concluding chapter.

Representations

Cognitive science is predicated on the belief that it is legitimate—in fact, necessary—to posit a separate level of analysis which can be called the "level of representation." When working at this level, a scientist traffics in such representational entities as symbols, rules, images—the stuff of representation which is found between input and output—and in addition, explores the ways in which these representational entities are joined, transformed, or contrasted with one another. This level is necessary in order to explain the variety of human behavior, action, and thought.

In opting for a representational level, the cognitive scientist is claiming that certain traditional ways of accounting for human thought are inade-

Cognitive Science: The First Decades

quate. The neuroscientist may choose to talk in terms of nerve cells, the historian or anthropologist in terms of cultural influences, the ordinary person or the writer of fiction in terms of the experiential or phenomenological level. While not questioning the utility of these levels for various purposes, the cognitive scientist rests his discipline on the assumption that, for scientific purposes, human cognitive activity must be described in terms of symbols, schemas, images, ideas, and other forms of mental representation.

In terms of ordinary language, it seems unremarkable to talk of human beings as having ideas, as forming images, as manipulating symbols, images, or languages in the mind. However, there is a huge gap between the use of such concepts in ordinary language and their elevation to the level of acceptable scientific constructs. Cautious theorists want to avoid positing elements or levels of explanation except when absolutely necessary; and they also want to be able to describe the structure and the mechanisms employed at a level before "going public" with its existence. While talk about the structure and mechanisms of the nervous system is relatively unproblematic—since its constituent units can (at least in principle) be seen and probed—agreement to talk of structure and processes at the level of mental representation has proved far more problematic.

Critics of the representational view are generally drawn from behaviorist ranks. Wielders of Ockham's razor, they believe that the construct of mind does more harm than good; that it makes more sense to talk about neurological structures or about overt behaviors, than about ideas, concepts, or rules; and that dwelling on a representational level is unnecessary, misleading, or incoherent.

Another line of criticism, less extreme but ultimately as crippling, accepts the need for common-sense talk about plans, intentions, beliefs, and the like but sees no need for a separate scientific language and level of analysis concerned with their mental representation: on this point of view, one should be able to go directly from plans to the nervous system, because it is *there*, ultimately, that all plans or intentions must be represented. Put in a formula, ordinary language plus neurology eliminate the need for talk of mental representations.

Of course, among scholars who accept the need for a level of representation, debates still rage. Indeed, contemporary theoretical talk among "card-carrying" cognitive scientists amounts, in a sense, to a discussion of the best ways of conceptualizing mental representations. Some investigators favor the view that there is but a single form of mental representation (usually, one that features propositions or statements); some believe in at least two forms of mental representation—one more like a picture (or image), the other closer to propositions; still others believe that it is possi-

ble to posit multiple forms of mental representation and that it is impossible to determine which is the correct one.

All cognitive scientists accept the truism that mental processes are ultimately represented in the central nervous system. But there is deep disagreement about the relevance of brain science to current work on cognition. Until recently, the majority viewpoint has held that cognitive science is best pursued apart from detailed knowledge of the nervous system—both because such knowledge has not yet been forthcoming and out of a desire to ensure the legitimacy of a separate level of mental representation. As the cognitive level becomes more secure, and as more discoveries are made in the brain sciences, this self-styled distancing may be reduced. Not surprisingly, neuroscientists (as a group) have shown the least enthusiasm for a representational account, whereas such an account is an article of faith among most psychologists, linguists, and computer scientists.

Computers

While not all cognitive scientists make the computer central to their daily work, nearly all have been strongly influenced by it. The computer serves, in the first place, as an "existence-proof": if a man-made machine can be said to reason, have goals, revise its behavior, transform information, and the like, human beings certainly deserve to be characterized in the same way. There is little doubt that the invention of computers in the 1930s and 1940s, and demonstrations of "thinking" in the computer in the 1950s, were powerfully liberating to scholars concerned with explaining the human mind.

In addition to serving as a model of human thought, the computer also serves as a valuable tool to cognitive scientific work: most cognitive scientists use it to analyze their data, and an increasing number attempt to simulate cognitive processes on it. Indeed, artificial intelligence, the science built around computer simulation, is considered by many the central discipline in cognitive science and the one most likely to crowd out, or render superfluous, other older fields of study.

In principle, it is possible to be a cognitive scientist without loving the computer; but in practice, skepticism about computers generally leads to skepticism about cognitive science. To some critics, computers are just the latest of a long series of inadequate models of human cognition (remember the switchboard, the hydraulic pump, or the hologram) and there is no reason to think that today's "buzz-model" will meet a happier fate. Viewing active organisms as "information-processing systems" seems a radical mistake to such critics. Computers are seen by others as mere playthings

Cognitive Science: The First Decades

which interfere with, rather than speed up, efforts to understand human thought. The fact that one can simulate any behavior in numerous ways may actually impede the search for the correct description of *human* behavior and thought. The excessive claims made by proponents of artificial intelligence are often quoted maliciously by those with little faith in man-made machines and programs.

Involvement with computers, and belief in their relevance as a model of human thought, is pervasive in cognitive science; but again, there are differences across disciplines. Intrinsic involvement with computers is a reliable gauge of the extent of a discipline's involvement with cognitive science. Computers are central in artificial intelligence, and only a few disgruntled computer scientists question the utility of the computer as a model for human cognition. In the fields of linguistics and psychology, one will encounter some reservations about a computational approach; and yet most practitioners of these disciplines do not bother to pick a feud with computerphiles.

When it comes to the remaining cognitive sciences, however, the relationship to the computer becomes increasingly problematic. Many anthropologists and many neuroscientists, irrespective of whether they happen to use computers in their own research, have yet to be convinced that the computer serves as a viable model of those aspects of cognition in which they are interested. Many neuroscientists feel that the brain will provide the answer in its own terms, without the need for an intervening computer model; many anthropologists feel that the key to human thought lies in historical and cultural forces that lie external to the human head and are difficult to conceptualize in computational terms. As for philosophers, their attitudes toward computers range from unabashed enthusiasm to virulent skepticism—which makes them a particularly interesting and important set of informants in any examination of cognitive science.

De-Emphasis on Affect, Context, Culture, and History

Though mainstream cognitive scientists do not necessarily bear any animus against the affective realm, against the context that surrounds any action or thought, or against historical or cultural analyses, in practice they attempt to factor out these elements to the maximum extent possible. So even do anthropologists when wearing their cognitive science hats. This may be a question of practicality: if one were to take into account these individualizing and phenomenalist elements, cognitive science might become impossible. In an effort to explain everything, one ends up explaining nothing. And so, at least provisionally, most cognitive scientists attempt

to so define and investigate problems that an adequate account can be given without resorting to these murky concepts.

Critics of cognitivism have responded in two principal ways. Some critics hold that factors like affect, history, or context will *never* be explicable by science: they are inherently humanistic or aesthetic dimensions, destined to fall within the province of other disciplines or practices. Since these factors are central to human experience, any science that attempts to exclude them is doomed from the start. Other critics agree that some or all of these features are of the essence in human experience, but do not feel that they are insusceptible to scientific explanation. Their quarrel with an antiseptic cognitive science is that it is wrong to bracket these dimensions artificially. Instead, cognitive scientists should from the first put their noses to the grindstone and incorporate such dimensions fully into their models of thought and behavior.

Belief in Interdisciplinary Studies

While there may eventually be a single cognitive science, all agree that it remains far off. Investigators drawn from a given discipline place their faith in productive interactions with practitioners from other disciplines: in the tradition of the Hixon and Macy symposiasts, they hope that, working together, they can achieve more powerful insights than have been obtained from the perspective of a single discipline. As examples, they point to current work in visual perception and in linguistic processing which has come to draw quite naturally on evidence from psychology, neuroscience, and artificial intelligence—so much so that disciplinary lines are beginning to blur.

Skeptics feel that you cannot make progress by compounding disciplines, and that it is more prudent to place each individual disciplinary house in order. Since it is also unclear *which* of the relevant disciplines will ultimately contribute to a cognitive science, and in which way, much valuable time may be wasted in ill-considered collaborations. From their vantage point, it is perfectly all right to have individual cognitive sciences but ill-considered to legislate a single seamless discipline. At most, there should be cooperation among disciplines—and never total fusion.

Rootedness in Classical Philosophical Problems

As already indicated, I consider classical philosophical problems to be a key ingredient in contemporary cognitive science and, in fact, find it difficult to conceive of cognitive science apart from them. The debates of the Greek philosophers, as well as of their successors in the Enlightenment,

Cognitive Science: The First Decades

stand out in many pages of cognitive scientific writing. I do not mean that these traditional questions have necessarily been phrased in the best way, or even that they can be answered, but rather that they serve as a logical point of departure for investigations in cognitive science.

In my discussions with cognitive scientists, however, I have found this precept to be contentious. Nor is it predictable which scientists, or which science, will agree with a philosophically based formulation of the new field. Some cognitive scientists from each discipline readily assent to the importance—indeed, the inevitability—of a philosophical grounding; while others find the whole philosophical enterprise of the past irrelevant to their concerns or even damaging to the cognitive scientific effort. We may well be dealing here with personal views about the utility of reading and debating classical authorities rather than with fundamental methodological aspects of cognitive science. But whatever the reason, cognitive scientists are scarcely of a single mind when it comes to the importance of the *Meno*, of Descartes's *Cogito*, or of Kant's *Critique*.

Precisely because the role of philosophy is controversial in the cognitive sciences, it is useful to explore the earlier history of philosophy. Only such a survey can prove that cognitive scientists—whether or not they are fully aware of it—are engaged in tackling those issues first identified by philosophers many decades or even many centuries ago. Scientists will differ on whether these questions were properly formulated, on whether philosophers made any significant progress in answering them, and on whether philosophers today have any proper role in a scientific enterprise. Indeed, even philosophers are divided on these issues. Still, it is worth reviewing their positions on these issues, for philosophers have, since classical times, taken as their special province the definition of human knowledge. Moreover, they have also pondered the nature and scope of the cognitive-scientific enterprise, and their conclusions merit serious examination.

In my own view, each of these symptoms or features of cognitive science were already discernible in the discussions of the 1940s and were widespread by the middle 1950s. A cognitive-science text will not necessarily exhibit or illustrate each of the symptoms, but few texts will be devoid of most of them. What legitimizes talk of cognitive science is the fact that these features were not in evidence a half-century ago; and to the extent that they once again pass from the scene, the era of cognitive science will be at an end.

Comments on the ultimate fate of cognitive science are most properly left to the conclusion of this study; but as a kind of guidepost to succeeding chapters, it may be useful to anticipate my principal conclusions. In my view, the initial intoxication with cognitive science was based on a shrewd

hunch: that human thought would turn out to resemble in significant respects the operations of the computer, and particularly the electronic serial digital computer which was becoming widespread in the middle of the century. It is still too early to say to what extent human thought processes are computational in this sense. Still, if I read the signs right, one of the chief results of the last few decades has been to call into question the extent to which higher human thought processes—those which we might consider most distinctively human—can be adequately approached in terms of this particular computational model.

Which leads to what I have termed the *computational paradox*. Paradoxically, the rigorous application of methods and models drawn from the computational realm has helped scientists to understand the ways in which human beings are not very much like these prototypical computers. This is not to say that no cognitive processes are computerlike—indeed, some very much resemble the computer. Even less is it to contend that cognitive processes cannot be modeled on a computer (after all, *anything* that can be clearly laid out can be so modeled). It is rather to report that the kind of systematic, logical, rational view of human cognition that pervaded the early literature of cognitive science does not adequately describe much of human thought and behavior. Cognitive science can still go on, but the question arises about whether one ought to remain on the lookout for more veridical models of human thought.

Even as cognitive science has spawned a paradox, it has also encountered a challenge. It seems clear from my investigation that mainstream cognitive science comfortably encompasses the disciplines of cognitive psychology, artificial intelligence, and large sections of philosophy and linguistics. But it seems equally clear that other disciplines mark a boundary for cognitive science. Much of neuroscience proceeds at a level of study where issues of representation and of the computer-as-model are not encountered. On the opposite end of the spectrum, much of anthropology has become disaffected with methods drawn from cognitive science, and there is a widespread (and possibly growing) belief that the issues most central to anthropology are better handled from a historical or a cultural or even a literary perspective.

And here inheres the challenge to cognitive science. It is important for cognitive science to establish its own autonomy and to demonstrate terrains in which computational and representational approaches are valid. I believe that cognitive science has already succeeded in this endeavor, though the scope of its enterprise may not be so wide as one would have wished.

If cognitive scientists want to give a complete account of the most central features of cognition, however, they (or other scientists) will have

to discover or construct the bridges connecting their discipline to neighboring areas of study—and, specifically, to neuroscience at the lower bound, so to speak, and to cultural studies at the upper. How to do this (or whether it can be done at all) is far from clear at this point; but unless the cognitive aspects of language or perception or problem solving can be joined to the neuroscientific and anthropological aspects, we will be left with a disembodied and incomplete discipline. Put differently, no one challenges the autonomy of biology, chemistry, and physics; but unless a single narrative can be woven from the components of atomic, molecular, and organic knowledge, the full nature of organic and inorganic matter will remain obscure.

All this risks getting ahead of our story, however. We have seen in the preceding pages how different factors present early in the century came together to form the bedrock of a new discipline. Ultimately, I want to take a close look at some of the best work in the discipline, so that I can properly evaluate its current status and its future prospects. To achieve this overview, however, it is necessary to consider how the very framing of questions within cognitive science grows out of philosophical writings of the past. By the same token, it is necessary to understand the particular histories, methods, and problems that have characterized the component cognitive sciences. Ultimately this philosophical and historical background has determined in large measure the nature and scope of current interdisciplinary cognitive-scientific efforts. In part II of this book, I shall take a careful look at the several disciplines whose existence made possible the idea of cognitive science and whose practitioners will determine the success of this enterprise.