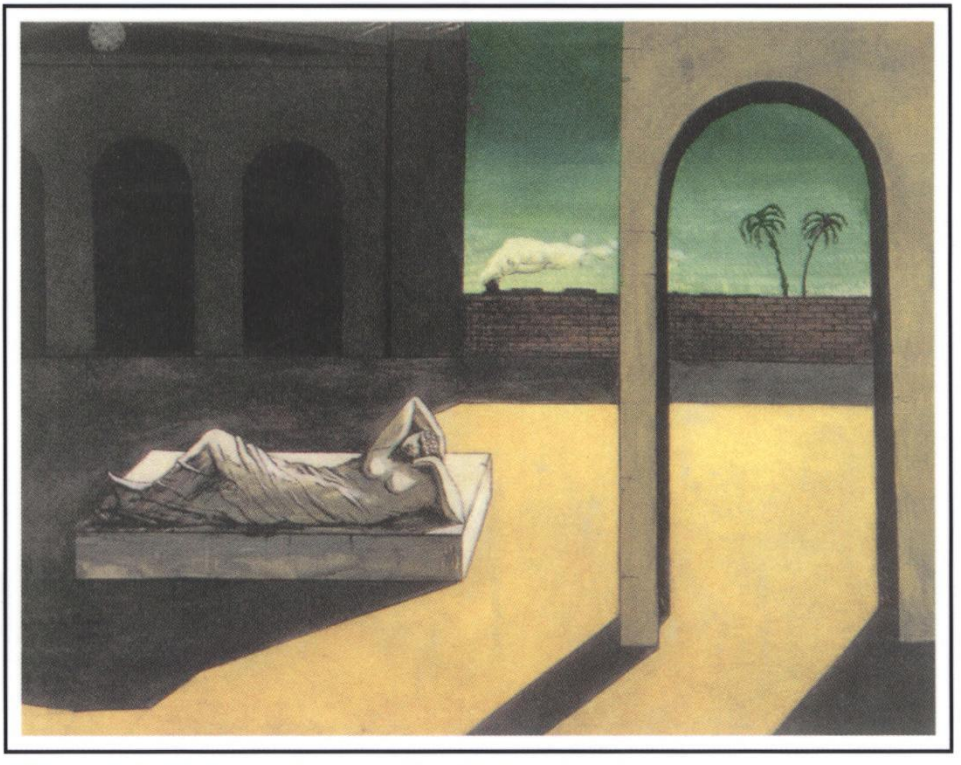
THE MUSE IN THE MACHINE

Computerizing the Poetry of Human Thought



David Gelernter

The MUSE in the MACHINE

Computerizing the Poetry of Human Thought

DAVID GELERNTER

THE FREE PRESS
A Division of Simon & Schuster
1230 Avenue of the Americas
New York, NY 10020

Manufactured in the United States of America

10 9 8 7 6 5 4 3 2 1

Library of Congress Cataloging-In-Publication Data

Gelernter, David Hillel.
The muse in the machine / David Gelernter.
p. cm
ISBN: 0-7432-3655-6
1. Artificial intelligence. 2. Consciousness. 3. Cgnition.
4. Emotions. I. Title.
Q335.G366 1994
006.3'01'9—dc20 93-49271
CIP

For information regarding special discounts for bulk purchases, please contact Simon & Schuster Special Sales at 1-800-456-6798 or business@simonandschuster.com

Contents

Acknowledgments

Acknowledgments are ordinarily a fairly cut-and-dried affair, but these are different. The summer of 1993, during which this book was completed, was the hardest of our lives. My wife and I needed lots of help, badly—and help was so generously forthcoming that I almost believe we will remember the summer more, in the end, for the kindness of family, friends, the community and total strangers than for the crime that occasioned it. If it in fact works out that way, isn't the only possible conclusion that good beats evil in the end?

Thanks are due first to our families—particularly to Cheri and Joel, to my parents and Jane's, Ed, Judy and other family members; and to our devoted dear friend Soledad Morales.

The manuscript would never have made it to the finish line without the extraordinary exertions of Nick Carriero and Chris Hatchell. I will always associate this book with their overwhelming generosity.

The Clemans, the Bialeks, Rabbi and Mrs. Benjamin Scolnic, the Agins, the Schwartzes, the Patkins and the Larrisons put themselves out repeatedly to help us with the children, which is the most important sort of help there is.

Among our friends and neighbors and my colleagues who helped

us in innumerable ways, large and small, the Lermans, Jane Milberg, Bob King, Susan Shapiro, Ed Keegan and Martin and Beverly Schultz were particularly indispensable. We deeply appreciate the help and good wishes of our friends Steve Nowick, the Arangos and the Jaggannathans, of Patricia MacDougall and the Yale Architecture Class of '83 (a truly great vintage!), of the Hamenachems, of the loads of people who brought us food, of the Yale police and particularly Officer Joe Vitale, of Ruth Anne Levinson and the JCC summer program staff, of Gail Brand and the whole Ezra community, Rabbi Weinberg and the B'nai Jacobites, President Levin, Provost Rodin and the Yale community *en masse*. I'm tempted to round out this huge list by appending the entire community of computer science researchers. Nothing moved and cheered me more last summer than the outpouring of good wishes that sloshed in daily from colleagues all over the world. People often write these notes because they feel they ought, without a sense that they will do much good; but they do immense good.

My graduate students carried on splendidly and cheered me up enormously. Thanks to all of them, and particularly Susanne Hupfer and Eric Freeman.

An academic is in bad shape without his books. Many thanks to Winston Atkins, Ellen Belcher and Paul Conway, whose expert ministration resuscitated so many of my tattered, burnt and drenched books and journals.

Last but (obviously) not least, my profound thanks to the medical people who so patiently reassembled me. They are a cast of thousands, too many to list. But the three whose exertions were most heroic are Rhonda Laidlaw, Marsha Dymarczyk and Dr. Henry Spinelli.

A few "normal" acknowledgments . . .

The research discussed here owes a great deal to my graduate students, but Scott Fertig in particular is responsible for much of the work and virtually all of the software that figures in this story. Dr. Mitch Sklar laid the basis years earlier with his own fine work.

My editor Susan Arellano did a great job with this quirky book. I am indebted to her, particularly for her patience.

Many parts of this book have gained from my conversations over the years with Professors Ivan Fox and Drew McDermott. (But

they are to be held wholly guiltless of any responsibility for the strange doctrines contained herein.)

Parts of Chapter 8 benefitted at an earlier stage from Barbara Harshav's superb editing and critical acumen. I'm only slightly ashamed to confess that if Mrs. Harshav liked this book and everyone else in the world hated it, I'd consider it a great success.

Every book I have written and every book I will write reflects, imperfectly, the wisdom of George McCorkle. He was an inspiring man.

And finally my wife Jane, from whom this summer I have learned the meaning of *eishet hayil. Ha-zorim b'dima, b'rina yiktsoru*.

Chapter One

A New Folk Psychology

It's hard to conceive offhand of a less promising consumer innovation than a computer that comes factory equipped with "emotions"—but here's a candidate: how about a "*spiritual*" computer? The spiritual computer spends its time pondering the mysteries of the universe, occasionally printing cryptic messages on its screen and otherwise ignoring the user altogether.

Here's what the "emotional" computer would do. You might describe a complicated medical case to it and ask for the diagnosis. The computer might give you a serious, telling answer, but add, ". . . still, I'm not happy with that; it doesn't feel right."

You might describe a complex legal case and ask for its advice. It's answered harder questions before, but on this occasion it might put you off with a comment about how the plaintiff reminds it of your sister.

You might describe an intricate stock deal and ask whether you should invest. In the past its advice has been solid—not infallible but better than any human's you've ever consulted, in part because it has billions of case histories down cold. But on this occasion it tells you, " 'Buzz off. I'm not in the mood. Let's talk about Jane Austen."

Who needs this kind of nonsense from a computer? Science does; in a broader sense we all do, because adding "emotions" to computers is the key to the biggest unsolved intellectual puzzle of our time: how thinking works. Oddly enough, our "emotional" computer will be capable of "spirituality" as well. No topic is further than spirituality from the interests of the researchers who are trying to understand the human mind. But as we will see, spirituality turns out to be central to cognitive psychology, and therefore to artificial intelligence, and therefore to computer science, and therefore to the whole history of science and technology.

Why these dramatic pronouncements and radical claims? Surely the science of mind is making fine progress without them.

Of course it is, but at the same time it suffers from a large and important problem. Here's an analogy.

Until 1759 sailors faced a big obstacle at sea: they only half knew where they were. Latitude was easily determined, but longitude was a mystery. Only when a meticulously accurate portable clock emerged at the end of a massive international research effort did the mystery of the longitude finally come clear (Landes 1983). To sailors, this development was a revelation; knowing half your position wasn't always a lot better than not knowing it at all.

Thought scientists today have a similar problem. The study of how thinking works is a big field—it encompasses philosophers of mind, cognitive psychologists, neurophysiologists and legions of frantically intense computer scientists bent on carrying off the greatest conjuring trick of all time, building minds out of computers. But thought science today is at sea. Despite monumental exertions, it has achieved a good grasp of no more than half the problem before it. Reasoning is one big part of human thought, and thought science has reasoning decently under control. Philosophers and psychologists understand it and computers, up to a point, can fake it. But there is one other big piece of the picture, which goes by many names: creativity, intuition, insight, metaphoric thinking, "holistic thinking"; all these tricks boil down at base to *drawing analogies*. Inventing a new analogy—hitching two thoughts together, sometimes two superficially *unrelated* thoughts—brings

about a new metaphor and, it is generally agreed, drives *creativity* as well. Studies (and intuition) suggest that creativity hinges on seeing an old problem in a new way, and this so-called "restructuring" process boils down at base to the discovery of new analogies. How analogical thinking works is the great unsolved problem, the unknowable longitude, of thought science. "It is striking that," as the philosopher Jerry Fodor remarks, "while everybody thinks analogical reasoning is an important ingredient in all sorts of cognitive achievements that we prize, nobody knows anything about how it works"—not even, Fodor adds (twisting the knife), in an "in the glass darkly sort of way" (1983, 107).

Thought scientists have exhausted themselves trying to solve the riddle. To cite only the most dramatic example, Roger Penrose (1989) posits by way of explanation an esoteric quantum mechanical freak of nature. Fodor (1983, 127) believes the answer is not merely unknown but unknowable—that such problems are just "bad candidates for scientific study."

I will present here a new model of human thought that puts analogy at the center of the action and offers what I claim is a direct, fairly simple explanation of how it works.

I'm about to tell a new and different story of human thought. But new and different does not mean concocted out of whole cloth. My method is more like the archeologist's who reconstructs an ancient vase out of its scattered shards. Most of my shards are important, intriguing, but (I claim) insufficiently appreciated results of experimental psychology. Some of my shards come from philosophy or computer science. Some—the most beautiful—are the work of poets, particularly the mind-obsessed English Romantics. A great poet has a more than average chance of possessing a mind that is, as William Wordsworth puts it, "haunted by herself." "In Egypt, Palestine, Greece and India the analysis of the mind had reached its noon and manhood," Samuel Taylor Coleridge writes in 1817, "while experimental research was still in its dawn and infancy" (Coleridge 1817/1975, 54). It's an overstatement, but an intriguing one. Modern thought science has no truck with poetry. But I am convinced that these poets can tell us deep, beautiful, *scientifically indispensable* facts about thought.

Like weathered pottery fragments, some of my pieces won't fit

exactly; others are missing and will require what appear (at first) to be big leaps of faith to reconstruct. But when everything is in place and a smooth, coherent shape emerges, those leaps will seem merely inevitable.

The result will be, in essence, a new "folk psychology"—cognitive science's slightly snide term for prescientific, commonsense psychology, the sort that can be done before even a single grant proposal gets funded. I will marshal a fair amount of scientific data to support my argument, but I will appeal more frequently to intuition and common sense. The results might or might not be convincing to scientists. I hope they are, but my main goal is to reach those readers for whom the human mind is not a profession but a passion. These aficionados are as likely to be poets, priests, gossips, or truck drivers as cognitive scientists. If I can show these people anything at all that broadens or deepens their grasp of this boundlessly fascinating, all-consuming topic even a little, I will be satisfied.

Dinner parties and boat rides

Here is my argument. Human thought is laid out in a continuous spectrum. Every human mind is a spectrum; every human mind possesses a broad continuous range of different ways in which to think. The way in which a person happens to be thinking at any given moment depends on a characteristic I'll call "mental focus." Focus can be high or low or medium; it changes throughout the day, not because the thinker consciously changes it, as he might consciously raise his arm, but in subliminal response to his physiological state as a whole. Fatigue (for example) makes focus go lower. Wide-awakeness makes it go higher.

Mental focus might sound like another way of saying "degree of alertness"; what's new is the way *cognition as a whole* changes in response to changing focus. High focus puts the thinker at the high end of the cognitive spectrum, and certain consequences follow. At the high end, thought is analytic and penetrating. It deals in abstractions and displays a "demythologized intelligence"—as the poet Robert Bly (1991) calls it—"that moves in a straight line made up of tiny bright links and is thereby dominated by linked facts." If a person's briefcase is stuck, he needs to open it and is in a high-

focus state, he will methodically run down the list of factors that cause briefcases to jam, plan a course of action and do it. His thoughts are well behaved. He has no doubt that they *are* (mere) thoughts; they do not impose themselves like hallucinations. Perceptions turn obediently into easily retrieved memories. In the future he will have no trouble recalling how he behaved during this particular stuck briefcase incident.

Almost all attempts to simulate thought on a computer have dealt exclusively with this narrow, high-focus band at the top of the spectrum.

As we set off down-spectrum, thinking becomes less penetrating and more diffuse, consciousness gradually "spreads out" and—this is a key point that I will spend much of the chapter explaining—emotion starts gradually to replace logical problem-solving as the glue of thought. The rest of the chapter pins down these vague pronouncements. When a briefcase jams and the owner's focus level is medium, instead of a cool logical analysis he is more likely to think "last time when I did this, it opened." Thought is less analytical, more concrete. He might simply give the thing a good whack. (Thought: a good whack usually helps out in situations like this.)

A few research efforts in artificial intelligence have attempted to reproduce this "medium-focus" type of thought—although, without noticing the spectrum itself.

Confronted with a stuck briefcase towards the bottom of the spectrum, the owner is likely not to solve the problem at all. He is more likely as he ponders the briefcase to find himself thinking "that was some hot muggy day when I bought this damned briefcase in Milford—did I overpay?—I pay more than I need to for most things. But I'm better than Bill Schwartz in that regard—Schwartz's dinner party last fall was sort of fun—Molly sure didn't want to go—she looked nice in that short midnight blue dress, though—Columbus Avenue, we got the thing in that shop around Seventy-sixth Street . . ."

Now suppose our thinker's focus is just a bit lower. His thought-stream might start off in the same way—might extend, say, through "Schwartz's dinner party last fall was sort of fun," and pause there for a while—various aspects of the party come to mind; and then, next thing he knows, the thinker might find himself contemplating

Long Island Sound as he crossed it on a ferry years ago, seated on the stern deck, admiring the glitter of the soft green water on a bright, hazy summer afternoon. That summer afternoon has no obvious connection of any kind with the Schwartz's dinner (which consisted, let's say, of an elaborate meal served on three end-to-end rickety card tables in a cramped SoHo apartment nowhere near Long Island Sound, and is memorable mainly for noise and the overwrought, grinning Schwartz's stories about the school board elections and Molly on her knees playing scrabble) but, for some reason or other, it comes to mind—

But this is a tremendously important phenomenon, this coming-to-mind seemingly out of thin air, when we are in a mental state that we might informally call "relaxed"—what I would call "low focus"—of thoughts bearing no evident relationship to their predecessors. Readers may doubt that it happens; just *because* it comes about in relaxed mental states only, the thinker rarely takes note of it, just passively experiences it. But psychologists have been aware of the phenomenon since at least 1823, when one Ludwig Börne (cited in Jones 1963, 160) wrote that, if you monitor your thinking uncritically for a few days, ". . . you will be amazed at what novel and startling thoughts have welled up in you." Modern studies that I will cite later back up this strange contention. Any reader who monitors his own thoughts for a while will discover the phenomenon for himself. Try it and you will see that it *does* happen.

Does it matter? Yes, enormously. These unexpected transitions from thought *A* to a seemingly unrelated thought *B* are (as I will discuss) exactly the occasions on which analogies are discovered and metaphors emerge. *A* and *B* don't *necessarily* make an analogy. The Schwartz affair and that trip across Long Island Sound might not be analogous in the least. But when analogies do emerge, they emerge in this way. It may be that, say, the Schwartz party and the ferry trip are two occasions on which our thinker had a wonderful time with a subtle undertone of anxiety, because of something that was going to happen the next day—

This mental leap from the noisy party to the placid boat ride is paradigmatic of the most significant unsolved problem of cognition. *Affect linking*, I will claim, is responsible for bringing these leaps about. They are not random (nor need they have anything to

do with repressed Freudian angst); they come about exactly when *two recollections engender the same emotion*, and they only happen towards the low-focus end of the spectrum.

Towards the lower end of the spectrum, affect linking causes creativity, metaphor, and in some cases spiritual mind-states to emerge. Other cognitive events accompany affect linking: thought grows ever more concrete. Recollection grows broader, more tangible and full of ambience and all-inclusive until eventually, a recollection becomes indistinguishable from a hallucination; and other things being equal, the illogic of dreaming waits at the bottom.

No computer program of which the author is aware attempts to simulate low-focus thought.

The cognitive spectrum provides us with a vantage point from which we can survey and make sense of human thought as a whole. More: it tells us something about the *dynamics* of human thought—its history over multiple time scales. Over the course of a day fatigue sets in and the character of thought changes. Over the first decades of life maturity comes about, and the character of thought changes. Over the millennia of human existence the modern mind gradually emerges, and the character of thought changes. I'll call these the Big Three cognitive transitions. They differ radically in character and take place over radically different time scales. But it's a curious fact that, if we view these three transitions from the spectrum's vantage point, they all three seem to tell the same underlying story. It is the story of gradual transition across the spectrum, from high focus towards low in the daily passage from awake to asleep; from low towards high in the development of a child or the emergence of the modern mind. The three transitions remain radically different, but the *underlying theme* turns out to be the same.

Now let me fill out this picture, make it more precise, and attempt to convince you that it is true.

The spectrum

Children have short attention spans.

There's nothing remarkable about that, but it perfectly epitomizes the *sort* of fact that appears to be wholly unconnected to top-

ics like computer science or the philosophy of mind. Most practitioners of those disciplines would agree: children have short attention spans . . . *so what*? In reality these short attention spans are connected to technology and philosophy by a dense network of facts in which they play one small but interesting part.

Everyone knows about children's short attention spans and, for the record, studies confirm them. Very young children have another related characteristic that's harder to study but just as interesting. If you listen to an intelligent two-and-a-half-year-old holding forth, you may notice that his conversation has a plastic quality: one topic turns abruptly into another, with no respect for narrative logic (Garbarino et al. 1989, 77, make observations along these lines). We might call this "stream of consciousness"-style conversation, but it strongly resembles another phenomenon as well. Dreams work this way. It's perfectly normal in the course of dreaming for one scene to transform itself abruptly and "illogically" into a different one. "The narrative organization of dreaming," Foulkes (1985, 35) circumspectly observes, "occasionally breaks down in a relatively spectacular way."

Now it's an interesting fact, less well known, that some ancient literature is marked by this same kind of plastic, illogically put-together quality. Take the Bible as an example. Most people are aware of the theories claiming that it consists of many separate narrative strands stitched together. Ask yourself: why would such a theory be plausible? Answer: *in part*, because of passages where there is no sustained logical thread. Where the narrative bobs and weaves and doesn't seem (*seem*) to make sense: like a young child, or a dream.

If you were a literary critic attempting to analyze a dream or the conversation of a two-year-old transcribed and presented to you as "literature," you'd notice all sorts of abrupt, implausible transitions, and many contradictions. You might easily decide that you were examining not the production of one author, but a bunch of separate accounts crudely patched together. In the case of the Bible there are other good reasons for guessing that there *are* multiple strands in play, and these good reasons have obscured a deeper point. Having observed a phenomenon and explained it we're

tempted, by and large, to go away satisfied. But suppose there is *more than one* explanation?

What is it about dreams, childhood, and the ancient mind that makes thought run, sometimes, in funny ways?

Could it be something they share? These *states of being* have been compared often before, for example in the mind-obsessed poetry of the English Romantics, or in Freud's anthropology. There's a certain obvious vulnerability in all three states. We sometimes dream about childhood; in the discredited "recapitulation" theory of earlier decades, the individual in the course of development was thought to retrace the evolutionary history of the species (Gould 1977). But I have something different in mind. I refer to a more specific resemblance, in *styles of thinking*, in the method by which thought-trains are assembled.

Returning to dreams: their defining quality is that, although they are constructed of memories, they seem real. "Dreams are hallucinations" (Hall 1966, 9). And on reflection, thoughts *about* childhood and *in* childhood are both tinged, also, with hallucinatory overtones.

Some authorities claim that young children actually do hallucinate routinely (for example Freud [1958], or Jaynes [1976]—of whom more shortly). Perhaps, but I have different and more subtle phenomena in mind.

First, certain childhood memories, suddenly recollected, can take an adult out of the present and place him in a different time and place with near-hallucinatory force. Not everyone knows the experience first-hand, but for those who do, it is striking. William Wordsworth explains: "Musing on them"—on memories of his early childhood—"often do I seem two consciousnesses, conscious of myself and of some other being" (*Prelude* I). This is a popular literary theme. E. B. White (1942): "I seemed to be living a dual existence." The Russian novelist Esther Salaman (1970), in a moving and beautiful study of what she calls "involuntary memories" of early childhood: "a *then* becomes a *now*." These are not hallucinations, but their vividness is uncanny. They reach out towards the hallucinatory.

As Salaman describes it, some thought or perception leads, for

no evident reason, to a sudden, vivid recollection of a childhood scene. If our low-focus briefcase thinker's contemplation of the Schwartz party had been followed by a sudden recollection, nearly hallucinatory in its intensity, of a mossy mountaintop home he had visited once at the age of six; if he had re-experienced the damp musty-smelling wooden-sided sandbox where he had passed an uneasy half hour alone—then he has experienced an involuntary memory of childhood.

Further: young children are often themselves observed to have vivid imaginations. (Again this is just obvious; but see, for example, Singer and Singer's 1990 discussion of imaginary playmates for one striking example.) And what does a "vivid imagination" mean? If I ask you to close your eyes and imagine lying on the beach, the better you succeed, the closer you've come to staging a small-scale auto-hallucination. If we say *you have a vivid imagination*, we mean that *what you imagine seems real to you*.

What is it about the mental states of dreaming and of early childhood that causes thought-trains to be built, sometimes, in funny ways, and *mere thoughts* to take on hallucinatory color? Is it something they share?

It's a striking fact that the same strange phenomenon, thought-processes tinged with the hallucinatory, is also associated with ancient man. The fact becomes clear when we react to ancient literature in the natural, unstifled way we would to a modern text. The psychologist Julian Jaynes read ???*The Iliad* this way, and was astonished (as any modern reader should be!) to discover that so many key decisions in the story are said to be dictated verbatim by the god on duty. We are accustomed to thinking of God or gods in sanitized, metaphoric terms. It's apt to completely pass us by when ancient man, speaking of the felt presence of the divine, *actually means it*! Jaynes concludes (1976) that ancient man was prone to routine auditory hallucinations—untenable, I think. Still, as judicious a critic as the great classicist E. R. Dodds observes that hallucinations were most likely "commoner in former times than they are today" (1951, 116). Another great classicist cautions that "we must not forget the power of hallucinations . . ." (Gilbert Murray 1955, 25); "primitive men seem to have dealt more freely than we

generally do with apparitions and voices and daemons of every kind" (26).

I would like to dispense with expert testimony and revert to common sense wherever I can. Hallucinations to the side, it is impossible to listen carefully to these texts and not to conclude that, in peak form, ancient man drew on a wildly vivid imagination. Wordsworth comments: "Turn where so'er I may,/ by night or day,/ the things which I could see I now can see no more" (*Ode: Intimations of Immortality*: 7–9). He refers to the lost, light-bathed vividness of childhood, but could just as easily have been describing the lost vividness of antiquity—when "holy were the haunted forest boughs,/ Holy the air, the water and the fire," as John Keats describes it (*Ode to Psyche*: 38–39). This vividly imagined, near tangible sense of holiness in the air *feels like* one step short of hallucination.

Now scholars have argued over whether ancient cognitive life was in fact different from our own. Some have claimed (for example, Miller 1986) that there is no reason to suppose that it was. Jaynes' view of Homeric man seems related to that of Snell (1953), which is pretty convincingly dismissed by Williams (1993). But there are concrete and specific reasons for believing that ancient thought *was* different: certain ancient texts prove that it was, by dint of making no sense under any other assumption. I will attempt to show in Chapter 8 that ancient thought-trains were just built differently, at least on some occasions, from our own; *and* we will see that these differently constructed thought-trains are associated with a vividly imagined or even hallucinatory mental life.

What is it that the mental states of dream, child, and ancient man share?

And why is it that, after the dream has passed, we have so much difficulty remembering it? And that after early childhood has passed, we have so much difficulty remembering *it*? And what's striking is that the *sort* of difficulty we have often seems to be very much the same in both cases. We feel intuitively that our failure to remember is often what psychologists call a "retrieval failure": the stuff is in there, but we just can't find it. Some psychologists suggest that this indeed is so, in the case of dreams (Cohen 1979) and of "childhood amnesia" (Shachtel 1947/1982).

Freud claimed that these two forgettings are related—that repression is the root of both. It's an intriguing claim that has occasioned many studies over the years, but they don't appear to bear it out, or at least not in any obvious way—although it is admittedly a very hard theory to test (Cohen 1979). The two kinds of forgetting *are* related, I think, but not in this way—or not only in this way.

No analogy between an individual and a social process can be more than rough, but if the analogy is evocative we are certainly within our rights at least to notice it. History itself emerges out of the mists, several millennia after the discovery of writing. The earliest stages of literate history are blanked out—except for the occasional vivid recollection. An event from the distant prehistoric past like the Biblical Flood recurs in many national mythologies. It haunts the world's literature like the "flashbulb memories" that fix a charged moment (say a president's assassination) in national consciousness (Brown and Kulik 1982, 23–40). Salaman's involuntary childhood memories are an even closer analogy. We can understand the Flood as mankind's involuntary childhood memory. It is a traumatic memory, as such memories (Salaman notes) so often are.

Why is it that memories of dream, childhood, and the ancient past are so hard to reclaim?

Symbolism is basic to dreaming. Everyone is aware of occasions on which *X* occurred in a dream, but we know immediately that *X* really meant *Y*. A symbol might be purely arbitrary. More often, there's some connection between "symbol" and "symbolized." In more ways than one, dream-thought traffics in *associations*, in unexpected connections. It's remarkable that childhood thought does too. "A common characteristic of young preschool children's conversation is *chaining*—that is, free-associating" (Garbarino et al. 1989, 78). On reflection, anyone who has ever listened seriously to a young child will agree; but unexpectedly, and at odds with the previous best guesses of developmental psychologists, a series of studies have shown young children to be strikingly good at *metaphor* as well (for example Gardner and Winner 1979, and Billow 1988). Metaphor hinges on meaningful but nonobvious connections.

What is it exactly that dream and child share?

Let's shift gears. What are children *bad* at?

Logic and analysis are not childhood specialties. More generally, children tend to be poor at manipulating abstractions. This is obvious and again, for the record, studies confirm it (e.g., Sternberg 1984). But adults are also conspicuously incompetent at logic, analysis and abstraction—when they are sleepy. A psychiatry textbook notes that drowsiness is characterized by, among other things, a "tendency to concrete thinking." (Freedman et al. 1976.) Adults are *very* bad at logical thought when they are asleep. A certain amount of problem-solving can occur in a dream, but dream-thought is notably *not* logical and not analytical. Nor is it abstract: dream imagery may be a concrete reduction or embodiment of a dreamer's thoughts. (Thus for example the psychologist C. S. Hall, who made exhaustive laboratory studies of dreaming, 1966, 95.) Freud (1900/1965, 330) remarked that words are treated in dreams as if they were concrete things; *concrete* is a word that keeps recurring in the literature on dreams.

Logic, analysis and abstraction are nonspecialties of ancient man as well. The great archeologist William Foxwell Albright (1940, 1964) used the term "proto-logical" to distinguish the thought style of the oldest class of western literature from ordinary logical thinking.

There is an obvious antithesis, then, between the mental universe of antiquity, dream and child and the analytic rigor of "rational man." But it isn't by a long shot merely the timeworn, threadbare antithesis between "metaphoric thought" and "logical thought." There is a whole rich world of mental phenomena that accompanies the nonrational states of mind. It is a strange-seeming world, uncanny even, in which thoughts are vivid (verging on hallucinatory) and thought-streams can be oddly and illogically built, where unexpected connections emerge like the features of a moonlit landscape as you grow used to the dark, and memories are slippery: it's hard to recall just what has happened, and where you have been . . .

That there are two styles of thinking is an old observation. It dates back to Aristotle. In modern psychology, an "analytic" or

"rational" or "convergent" style is contrasted with something called "divergent" or "productive" or "lateral" or "primary process" or "metaphoric" thought. Sometimes "left-brain thinking" is held up to "right brain" in a broadly similar way. None of these dichotomies is a good match to the antithesis I've been describing. But in any case, the dichotomy itself is not the point. Mine is a story of *continuity*: of one cognitive style unfolding smoothly into another. What the dichotomists miss is merely the crux of the matter: that these two styles are *connected*; that a spectrum joins them, a spectrum that runs in one continuous, subtly graded arc, from the intense violet of logical analysis all the way downward into the soft slow red of sleep. They fail to grasp that there *are* any colors but red and violet; most important, that red *is* violet!—but for a *single* adjustment. Merely change the frequency. Leave everything else the same.

Let's invent something called a person's mental *focus*. For now I won't say what it is—it doesn't matter—beyond asserting that it correlates with current position on the mental spectrum. One end of the spectrum corresponds to maximum mental focus, and the other end to minimum. As we twiddle the knob from the high "analytic" end to the low "metaphoric" side, we are gradually turning *mental focus* lower.

Many separate cognitive events accompany the move down-spectrum, ranging from a loss of control over the thought-stream to an increased propensity to have creative insights and to encounter vivid imaginings or even hallucinations, a relaxation of logic, a loss of "goal directedness" and *all* directedness, and the emergence of emotion as the main glue of thought. But it is essential to grasp that this is no mere jumble of processes that just happen (allegedly) to be correlated. *These are all consequences of one underlying event.* They are all consequences of the *relaxation* or *lowering* or *widening* of the crucial property I have called mental focus. Lowering mental focus causes all of these processes to occur.

This cognitive spectrum, which is merely lots of well-known observations connected in a new way, is also the central fact of human thought. And it is the prime mover of cognitive history: of the cognitive transitions occurring over a day, over a lifetime and over the historical emergence of rational man from prelogical

antiquity. A broad range of other hitherto scattered facts will fit comfortably into this framework.

The existence of the cognitive spectrum is supported, at least indirectly, by a fair amount of data, which I will be marshaling throughout the book. And although the spectrum itself is novel, many other thinkers have laid out theories that are similar in part or sympathetic to the spirit of the whole. Nonetheless: the spectrum idea is radically at odds with today's mainstream thought science. The mainstream sees thought as reasoning, problem solving, analysis and mental modeling, with an occasional dash of common sense. When the cognitive psychologist Robert Sternberg (1982, 225) writes that "reasoning, problem solving, and intelligence are so closely interrelated that it is often difficult to tell them apart," he is trying to tell us something: that intelligent thought is directed at a goal, dedicated to the rational solution of a problem. Anyone who has ever looked out the window, let his mind wander and realized that indeed he is still, right now, manifesting his intelligence, knows that this is false. From my perspective, the mainstream's view of thought is a parody. Thinking is vastly richer than that. A thought theory that never comes to grips with intuition, hallucination, spirituality or dreaming cannot possibly be a serious account of cognition.

Let's methodically descend the cognitive spectrum. What happens when, starting in a high-focus, alert, analytic frame of mind, we gradually turn focus lower?

Driving is an odd but useful starting point. It's possible to drive in a state of high focus, intent and alert. Beginning drivers have no choice, and everyone ticks into this style occasionally. But ordinarily, a practiced driver allows a great deal of focus to boil off: doesn't concentrate intently on driving; devotes to driving a "lower focus" mental effort. The result is that he finds himself thinking about other things as he drives. Dennett points out (1991) that we shouldn't imagine driving to be carried out "unconsciously" under these circumstances, that it's merely absorbing a lowered level of attention.

As focus falls, attention becomes more diffuse. No single point

in the perceptual or mental landscape is sharply illuminated. Many points are less strongly lit. (But this is no "spotlight of consciousness" theory, of the sort criticized by Allport 1989, 631–682. His "spotlight" is a fixed-focus beam playing across memory. It will become clear, if it isn't already, that the model and intent here are completely different.)

When focus drains away to disperse over a wider mental landscape, a person has less attention to devote to driving, and more to something else. That "something else" can't *itself* be a high-focus activity: a certain amount of mental focus is "pinned down" by driving or whatever. As a person moves down-spectrum, focus gets dispersed—meaning what, exactly? Here's a first clue. This lowered-focus state is specifically associated with another crucial mental phenomenon, *creativity*. It is consistently singled out as *the* mental state that gives rise to creativity or inspiration.

Many writers have said so. For example, Roger Penrose. Creative thoughts are most likely to occur to him when he is thinking about a problem "perhaps vaguely," "consciously, but maybe at a low level just at the back of my mind. It might well be that I am engaged in some other rather relaxing activity; shaving would be a good example" (Penrose 1989, 419). Shaving, driving, whatever—this is strikingly similar to the (in our terms) "lowered focus" state that Dennett describes.

Creativity in turn is a phenomenon that centers, according to a fairly solid consensus, on *restructuring* (Holyoak 1990, 117–146): on finding unexpected *analogies* that allow a person to see old problems in a new light. If lowering your mental focus allows creativity to happen, then it must encourage the discovery of unexpected connections. Analogies, metaphors, unexpected connections start to emerge as we make our way down-spectrum.

Unexpected connections are also part of childhood cognition ("chaining," metaphor); and are central to dreaming.

One special facet of creativity is the next clue, to another related phenomenon that intensifies as we continue down-spectrum. We say that an inspiration "hits us," that it happens unsought. An inspiration seems to happen *to* us, not to be a willed consequence of our own mental acts. Many witnesses have said so—that inspirations "come suddenly, as we say, 'into a man's head' " Dodds (1951, 11).

Salaman (1970, 59) in a striking observation links her idea of *involuntary* memory directly to creativity: "An involuntary memory has this in common with the solution of an artistic, mathematical or any other problem: it is a swift and usually unexpected contribution of the unconscious mind. . . ." By *unconscious mind* she means that an unexpected link *just happens*, unsought (or, not consciously willed).

Of course, to the extent thoughts happen *to you*, your thought-stream is out of control. An inspiration or a creative thought is only a minor, momentary loss of control. "Fancy," after all, is "blended and modified by choice" according the great romantic poet Samuel Taylor Coleridge (in Edel 1982, 14). Creative thoughts hit you unexpectedly, but then you pick yourself up, dust yourself off and get on with it, regaining control over your thought-stream and putting your new insight to work. Psychoanalysts cultivate this medium level of focus—asking to be knocked over by an association or an insight, but primed to leap immediately to their feet again and retake control (see, e.g., Gay 1985, 42).

But as a person ventures further down-spectrum, his grip loosens and his train of thought runs away from him, and he comes closer to being swept away. A single creative fancy is a momentary loss of control; an involuntary memory is a step further, towards the outer limits of hallucination. Then there is such a thing as a "creative *state*" in which you are closer to a sustained loss of control—where a flood of ideas and not just one seizes your attention. Thus Dodds again, quoting E. M. Forster: "In the creative state a man is taken out of himself." (64). Creative states presumably do not happen to everybody. But suppose, for the sake of argument, that our briefcase thinker is a talented architect; suppose that, in the course of his ferry trip across Long Island Sound, idle contemplation of the ferry's wheelhouse leads him suddenly to conceive the concert hall he is designing in vertical rather than horizontal terms. So he finds a sheet of paper and sketches for an hour, more or less oblivious of his surroundings, as ideas tumble out and the entire design reorganizes itself. That's a creative state.

Taken out of himself: remember that we are still talking about mere sustained creativity. The poet Percy Bysshe Shelley makes us feel that a sustained *spiritual* state of mind might be even closer to a complete loss of control over your own thought-train:

antiquity are big, complicated stories reflecting countless influences, but they all share one underlying theme: the slow, steady trek of the cognitive campsite up or (for sleep) down the cognitive spectrum. And by the way, why is the sound of rain restful? *For the same reason that shaving is "inspirational"—*

Driving, shaving, whatever, are excuses to diffuse consciousness, *mental focus*, over a wider field. They pin down *some* mental focus but to the extent they are uninteresting (that is, unpleasurable) in the role of sole thought-focus, they encourage *some* mental focus to wander elsewhere. Searle quotes William James' formulation that "consciousness goes away from where it is not needed," but points out that "attention goes away from where it is not needed" is closer to the truth (1992, 139). Of course, as Searle concedes, it doesn't go away *entirely*. So long as a person is devoting a certain amount of attention to the wheel or the razor, the other ongoing mental activity *cannot* be a high-focus activity.

Low-focus mental activity sets the stage in turn for creativity, and at the extreme for sleep. No matter how physically exhausted a person may be, he cannot achieve normal sleep unless he is in a low-focus mental state. (I'll argue that insomnia is in many cases nothing more or less than unwanted high-focus thought.) The sound of rain or anything similar helps to lower mental focus, by pinning down *some part* of the thinker's attention, thus leaving less to be shared by any other mental activities. By insinuating itself into awareness, the sound of rain broadens and thus lowers focus. A louder or more "interesting" sound wouldn't have the same effect—just as hard driving or a dull blade doesn't lend itself to restful, productive musing: it steals the thinker's whole attention and *itself* becomes a target for high-focus thought. But allowing the sound of rain merely *to enter your awareness* drives focus wider, and in so doing forces thinking down-spectrum, towards sleep.

So we have a notion of "focus," and twiddling the focus knob controls your location on the mental spectrum. It's time to expand the model one small step further.

Thought at high focus isn't merely narrow; it's penetrating. To associate "penetrating" with "narrow" seems natural, suggesting

that, other things being equal, the same force spread over a smaller area will have a more concentrated effect. So: the high-focus thinker can spindle a stack of memories—glom them together into a temporary composite (a "memory sandwich") and examine *one* aspect of *all* of them. High-focus thought is capable of penetrating a whole stack of memories at once. When our high-focus briefcase thinker notes that the combination lock won't open, and it occurs to him that such locks may have had their combinations reset—that hypothesis results, probably, from many experiences with combination locks, not from the recollection of a particular incident. Or suppose the same morning he'd noticed it was snowing, and grabbed his scarf. He happens to know that if it is snowing, it must be cold. How does he know? On countless occasions reaching back to early childhood, he's observed that snowy days are cold. At some point the child stacked up many memories of snowy days, noticed that, for all their idiosyncracies, *cold* figured in all of them—and has known ever since that snowy days are cold.

As the focus dial is nudged lower, attention spreads out and becomes not *narrow and penetrating* but *broad, inclusive* and "superficial"—*many aspects* of *one* memory hold the thinker's attention, not *one* aspect of *many*.

When a memory is recalled at low focus, it is *the whole memory* that comes to mind. Not merely the technique whereby the briefcase was successfully opened when it was last jammed (if opening a jammed briefcase is the task at hand), but the *whole memory* in which this briefcase-opening experience is embedded: the jet fuel and stale cigarette smell of the airport waiting lounge, the headache, the destination, the weather outside and anything else that memory succeeded in trapping on that occasion.

"Abstract" versus "concrete"—think what the two words imply, and what we should conclude on that basis about the mental states that dwell on each. An abstraction is defined by the *few* traits shared by *many* instances. What is *concrete* includes many traits of one instance. Think of bravery in the abstract versus a concrete instance. The concrete instance is full of particular, individual details that have nothing to do with the abstract concept—

Now connect these obvious facts to low-focus thought in general and to dreaming as its radical end point. We know that dream-thought tends to be concrete. We also know that it is "out of control" in several ways. It is hallucinatory, and it is capable of presenting to the dreamer thoughts and memories that he doesn't want to have—that he has rejected or (perhaps) repressed in waking life. "Impulses that are kept in check during waking life" emerge in dreams (Hall 1966, 14). The reality of concrete thinking may be associated with both aspects of lost mental control. Concrete thought has the richness of texture that is required if we are to be enfolded by an imagined or hallucinated world. By exactly the same token, concrete thought is *inclusive*. We will explore later the connection between this all-inclusiveness, this inability to focus and thus to edit, and the occasionally tactless mental landscape of dreams. The power to focus-on is also the power to turn-away-from.

But there is still one other facet of the loss of mental control at low focus: the way thoughts "hit you," occur "by themselves"; the way a thought-stream can run away from the thinker. I'll return to it.

Of course abstract thinking emerges at high focus, and low-focus thought is concrete. An abstraction is merely a bunch of related memories "held up to the light" and examined as a single bundle. The salient features of this memory bundle are exactly the features that *define the abstraction* (the abstraction *blue, dog, true, irony*) as opposed to any one individual blue thing, dog, truth, or ironic incident. Because high-focus thought is penetrating and picks out details, it's the state of mind in which shared details come to light.

And because abstraction is a high-focus phenomenon, so is language, insofar as language deals with the meaning of words. A child learns what "blue" means by forming an abstraction. *Blue* is simply what all things called "blue" have in common. "Children under the age of seven are more responsive and attuned to nonverbal language—gestures and actions—than they are to words. They interpret words by relying on their understanding of the concrete circumstances around them" (Garbarino et al. 1989, 68). The emergence of language in childhood is a complicated process, but it takes place to a steady underlying beat: children as they mature

move their cognitive camping grounds gradually, inexorably up-spectrum—and gradually, inexorably, language becomes possible.

Memories *recalled at high focus* are the stuff of abstractions. A thinker pierces them *en masse* with his focussed attention and picks out the salient shared details. What about memories *formed* at high focus? Perceptions becomes memories—"turn imperceptibly into memories" (Dennett 1991, 136); it follows that, other things being equal, *memories formed at high focus are easily recalled*. A thinker pulls a memory back into awareness by grabbing some detail that stands out. At high focus, details do stand out. They become natural handles for pulling the scene back into awareness.

At low focus, the mind is occupied by the whole of *one* scene. Hence, thought tends to be concrete—that is, full of the particular, specific, idiosyncratic detail that sets a scene and creates an ambience. (In drowsiness, your thought grows "concrete." Children's thinking tends to the "concrete.") The breadth and inclusiveness of memories *recalled at low focus* give those recollections their potentially hallucinatory power to transport the thinker to another time and place. The breadth and inclusiveness of memories *formed at low focus* make them slippery. What emerges isn't detail but ambience. If a memory is formed, not of anything in particular but of everything in general, then, other things being equal, *it's hard to grasp such a memory*; hard to find any mental handhold. (The very quirk that makes memories slippery is the same quirk, of course, that makes thought vivid and potentially hallucinatory; and makes thinking not abstract but concrete.)

Childhood memories are hard to recall. Dreams are hard to recall. Many factors are at work in each case. In no instance do I wish to wave a magic wand over a complex phenomenon and announce that low-focus thought explains it all. But low-focus thought is an ingredient: it *sets the stage*. It creates the necessary condition. Childhood and dreaming are rich in low-focus thought—and that is a recipe for a vivid mental life, and amnesia afterwards. Young children themselves are bad at "deliberate recall," at reconstructing past events on demand, but they "often display remarkably accurate recognition memory" (Garbarino et al. 1989, 43)—when a *whole scene* is presented to them for examina-

tion, they can tell you whether or not they have encountered it before.

We've made progress, but are still a long way from a complete picture of the spectrum. The scheme as I've described it has a problem: as we move down spectrum, thought trains *ought* to unravel, simply come apart at the seams.

After all, the stuff that is conventionally held to be the prime glue of thought—logic, analysis, problem solving—fades as we move down spectrum and eventually vanishes. So what is it that holds thought together? Why doesn't your thought-train behave, as focus sinks, like a neatly stacked tower of blocks that goes up in a spaceship? As gravity slackens, the tower gradually becomes prone to dissolve into disconnected floating objects. But thought holds together. Some "mystery force" must kick in as logic fades. A force *different* from logic but a force nonetheless. What is that force?

To put the question another way: why should a thing like shaving induce a loss of control over your own thought-stream?

We have seen that activities like shaving, which may lower your mental focus, are associated with creativity. We know that creativity is felt by folk psychologists to be (as Dodds says) "not the work of the *ego*." What is it about a thing like shaving that might cause "a man to be taken out of himself"?

There is a basic mechanism that governs the assembly of thought-trains. It is a regular process, the cognitive equivalent of a pulse. On the basis of your "current thought" (say, "the briefcase is jammed"), you present a probe or cue to memory ("jammed briefcase"); memories emerge in response (involving, let's say, jammed briefcases, other jammed things, other kinds of briefcase); those memories are the basis of a *new* current thought ("usually if I do this, it opens").

This cycle is fundamental. It is mental breathing, or a cognitive heartbeat: a current thought giving rise to a probe, probe to memories, memories to a *new* current thought, new current thought to another probe . . . It's a simple observation, not so different at base from the old associationist doctrines of the British empiricists.

The "pulse" analogy is loose: the *cognitive* beat can be interrupt-

> I hold, that association depends in a much greater degree on the recurrence of resembling states of Feeling, than on Trains of Ideas . . . I almost think, that Ideas *never* recall Ideas, as far as they are Ideas—any more than Leaves in a forest create each other's motion—The Breeze it is that runs through them; it is the Soul, the state of Feeling—(cited in Willey 1957/1973, 96)

Wordsworth writes, in *Tintern Abbey*, of *a serene and blessed mood* in which *the affections gently lead us on . . .*

Affections are the key. To find the gist of low-focus thought, look to the *emotions*.

Virtually any serious survey of the topic will include a sentence such as, "The area of emotional experience and behavior is one of the most confused and ill-defined in psychology" (Izard, cited in Ekman 1977, 9). But we have to start somewhere. Lazarus and Smith cite the following from Drever's 1964 *Dictionary of Psychology*:

> *Emotion*: differently described and explained by different psychologists, but all agree that it is a complex state of the organism, involving bodily changes of a widespread character—in breathing, pulse, gland secretion etc.—and, on the mental side, a state of excitement or perturbation . . . (610)

Let's accept this definition, but turn the volume down. An emotion is a mental state with physical correlates; it is a *felt* state of mind, where "felt" means that signals reach the brain that are interpreted as bodily sensations, however fleeting and subtle.

"Happy" and "sad" and "angry" and "scared" are emotions, but it's crucial that the list not stop there—that it also encompass what are sometimes called the "aesthetic" emotions. Happy, sad, and so on are the emotional equivalent of saturated primary colors. But there is a vast range of subtler intermediates.

It's hard to get a purchase on such mental states—what a person might feel on an unexpectedly warm spring morning, on an empty beach in winter, on driving past a childhood home, or hearing a clock tick loudly in the silence, or pounding a nail squarely into a wooden plank. Such occasions might evoke an emotional response.