Human Cognitive Evolution: What We Were, What We Are Becoming

BY MERLIN DONALD

CULTURES ARE MADE UP of minds; they are by definition the products of individual minds in interaction, and therefore the form of the individual mind constrains the type of culture that any given species will produce. This is widely acknowledged in ethology, where the cognitive capacities of a given species are viewed as direct products of evolutionary forces acting simultaneously on the social and cognitive levels. The patterns of animal group behavior that some ethologists call "cultures" are direct reflections of the cognitive makeup of a species; for instance, the very different cognitive sensitivities of dogs and apes are precisely tailored to their unique patterns of social behavior, and vice versa. So must it be with humans; our special mental features shape our social interactions and cultural artifacts. And as our peculiar type of mind evolved, our cultures changed, in fundamental ways, from the apelike cultures that preceded them.

Bridging the Gap Between Presymbolic and Symbolic Representation

Humans are the only minds in nature that have invented symbols. Some apes can be trained to use human symbols in a limited way, but they have never *invented* them in the wild; and most animals cannot even be taught to use them. Humans, on the other hand, created the first symbols from whole cloth. Crossing the abyss between presymbolic cognition and symbol-driven cognition was a uniquely human adventure, and consequently there is a huge gap between human culture and the rest of the animal kingdom. Any comprehensive theory of human cognitive evolution ultimately boils down to a hypothesis about how this gap came to be bridged.

The importance of this gap to the cognitive sciences cannot be overstated. Cognitive science is broadly divided between the artificial intelligence or AI tradition, which builds symbol-driven models of mind, and the "neural net" tradition, which develops models of simulated nervous systems that learn without using symbols, by building hologram-like memories of experience. A neural net is basically a diffuse *tabula rasa* network of randomly interconnected memory units which learns from environmental feedback. In other words, the neural net learns much the way many animals are presumed to, by building up associations in a relatively unstructured memory network, in a manner the behaviorists would have approved. Existing artificial neural nets are not yet very powerful; they might approximate the capabilities of some of the simpler insect nervous systems, and it is not yet clear how far up the evolutionary scale these modeling principles can be carried. Nevertheless, neural nets have a degree of independent knowledge of the world.

AI models, on the other hand, depend on having preordained symbolic tools given to them by a programmer, and these symbols—in the form of elementary categories and rules—are used to construct symbolic descriptions of the world, rather like those that humans build with language. But there is a crucial difference between such artificial "expert systems" and the human mind. Expert systems have no independent knowledge of the world, and remain locked in at the symbolic level, so that to understand a sentence, they are essentially limited to looking up symbols in a kind of computational dictionary, each definition pointing on to more words or actions, in endless circles of dictionary entries. In such a system there is no way out, no path back to a model of the "real world," and symbols can be understood only in terms of other symbols.

Since, as Wittgenstein observed, the vast majority of words cannot be adequately defined with other words, this is not a trivial limitation. The AI tradition has run into a brick wall in its development, as Dreyfus predicted twenty years ago,[1] precisely because it cannot cross the invisible boundary line between presymbolic and symbolic representation and access the holistic, nonsymbolic kinds of knowledge that humans use to inform their symbolic constructs.[2] As a result, expert systems remain heavily dependent on the human user. Expert systems are at the cutting edge of evolution in one sense, because they extend the human nervous system out into the world and are therefore situated at the extreme opposite end of the evolutionary continuum from the neural net; but they possess little or no independent intelligence.

The key question of human cognitive evolution might be rephrased in terms of this dichotomy: somewhere in human evolution the evolving mammalian nervous system must have acquired the mechanisms needed for symbol-based thought while retaining its original knowledge base. To extend the metaphor (recognizing that it is a metaphor), it is as if the evolving mammalian mind enriched its archaic neural net strategy by inventing various symbol-based devices for representing reality. This is presumably why the human brain doesn't suffer from the present limitations of AI; it has kept

[1] H. L. Dreyfus, *What Computers Can't Do: A Critique of Artificial Reason* (New York: Harper & Row, 1972).

[2] There is an excellent discussion of this problem in a series of papers published in Z. Pylyshyn, ed., *The Robot's Dilemma: The Frame Problem in Artificial Intelligence* (Norwood, N.J.: Ablex, 1987).

the basic primate knowledge systems, while inventing more powerful ones, initially to serve some nonsymbolic representational agenda. This focuses the evolutionary question somewhat: how could the evolving primate nervous systems of early hominids have crossed the presymbolic gap? And what are the necessary cognitive antecedents of symbolic invention?

Gaining Access to the Contents of Memory

It is a difficult concept for humans to grasp, but animals cannot gain voluntary access to their own memory banks. This is because, like neural nets, they depend on the environment for memory access; in effect, they can only react to whatever situations the environment presents to them. Thus they are creatures of conditioning, and cannot "think" except in terms of reacting to the present or immediately past environment (this includes even their use of trainer-provided symbols, which is very concrete). We tend to anthropomorphize our interpretations of animal behavior, and assume they are also able to recall things voluntarily. But there is no evidence that they can do this; humans alone have self-initiated access to memory, or what may be called "autocuing."[3]

Autocuing is the ability to voluntarily recall specific memory items independently of the environment. Consider an animal moving through a forest; its behavior is defined by the external environment, and it can be very clever indeed in dealing with that environment. But humans can move through the same forest "thinking about" something totally unrelated to the immediate environment—for instance, the recent election, or a movie, or an item in the newspaper. In "thinking about" some topic the thinker pulls an item out of memory, reflects on it, accesses another memory item, connects this to the previous idea, and so

[3] Free-associative memory phenomena like dreams or daydreams are not evidence for true autocuing, because they are not voluntary or intentional.

on, in recurrent loops. This reflective skill depends on voluntary autocuing; each memory item has to be sought out, precisely located, and retrieved, preferably without retrieving a batch of other unwanted items, and without relying on the environment to come up with the relevant cues to help find the item. Our ability to transcend the immediate environment could not have developed without autocuing skill.[4]

In terms of the neural net/AI dichotomy raised above, the first symbolic memory representations had to gain explicit access to the implicit knowledge latent in neural nets.[5] The initial adaptive value of the first representational inventions of early humans would have been their ability to provide retrieval paths to a knowledge base that was already present, but not voluntarily accessible, in the primate brain. But, given the functional arrangement of the primate brain, where would such paths have been built?

Evolutionary Landmarks and Transition Periods

Human cognitive evolution was a multidimensional process spread out over a period of at least 2 million years, and we may never know the story in fine detail unless there are striking breakthroughs in fields like behavior genetics or the analysis of fossil DNA. But a reflection on archaeological reconstructions of our emergence might at least provide a framework in which

[4] Note that I am not saying we can introspect on the process by which we voluntarily access memory. We do not have to be aware of the retrieval process to have voluntary control over it: language is "voluntary" cognition, but we have no awareness of where the words are coming from when we speak.

[5] P. Graf and D. L. Schacter, "Implicit and Explicit Memory for New Associations in Normal and Amnesic Subjects," *Journal of Experimental Psychology: Learning, Memory and Cognition* 11 (1985): 501–518. The authors distinguished "implicit" from "explicit" memory, and cited evidence that explicit memory has unique neurological underpinnings (amnesics typically lose explicit memory while retaining implicit memory). Using their terminology, we could say that human cognitive evolution is largely about acquiring the retrieval architecture needed for explicit memory.

to phrase the right questions. Two major cognitive landmarks stand out: the era of *Homo erectus* (approximately 1.5 million to 300,000 years before present) and the era of archaic *Homo sapiens* (roughly 500,000 to 100,000 years before present). Both periods were marked by anatomical and cultural signs of major cognitive change.

The First Cognitive Breakthrough: Pure Mimesis. There were very major changes in the human genome approximately 1.5 million years ago, marked by the appearance of a new hominid species, *Homo erectus*, whose brain eventually reached over 70 percent of our modern brain-to-body ratio, and whose cultural achievements obviously required some form of improved memory capacity. This species produced (and used) sophisticated stone tools, devised long-distance hunting strategies, including the construction of seasonal base camps, and migrated out of Africa over much of the Eurasian landmass, adapting to a wide variety of environments. Importantly, these ancient hominids improved significantly over time, eventually achieving the continuous use of fire. Could they have achieved this without language?

Many evolutionary theorists don't want to consider this possibility; they are fixed on the idea that there was only one great cognitive breakthrough, language, that this breakthrough came early, with *Homo erectus*, and that all higher human mental abilities followed from this breakthrough. Parker and Gibson[6] and more recently Bickerton[7] have argued that some form of "protolanguage," perhaps resembling the agrammatical speech of a two-year-old child, must have existed at the time of *Homo erectus*. This would explain early hominid cultural achievements with a single adaptation—a sort of grammarless language—that later evolved into modern speech capacity, when a capacity for grammar was added.

[6] S. T. Parker and K. R. Gibson, "A Developmental Model for the Evolution of Language and Intelligence in Early Hominids," *Behavioral and Brain Sciences* 2 (1979): 367–408.

[7] D. Bickerton, *Language and Species* (Chicago: University of Chicago Press, 1990).

I find this idea unconvincing. First, archaeological evidence doesn't place speech so early in evolution; neither of the principal markers for human language—the descended larynx and rapid cultural change—appears in the archaeological record until *Homo sapiens*, who evolved over a million years later. Second, early hominids had no existing linguistic environment, and therefore even the simplest protolanguage would have required some capacity for lexical invention. This issue is crucial, because it raises the issue of the autocuing of memory: lexical inventions must be autocuable. True linguistic symbols, even the simplest, could not suddenly pop up in evolution before first having established some principle of voluntary memory retrieval in the hominid brain; to be at all useful, lexical inventions had to be voluntarily retrievable as well as modifiable. They also had to be truly representational acts, intentionally modeling some aspect of reality.[8]

The first step, therefore, before lexical invention became a realistic possibility was establishing voluntary retrieval, or autocuing, in the prelinguistic brain. The same adaptation would ideally also have provided the cognitive prerequisite for a number of human *nonverbal* representational skills; language is not the only uniquely human cognitive advantage that has to be explained in our evolution.[9] If all of our higher thought-skills were based on our linguistic capacity, how could we account for the virtual autonomy of some nonverbal forms

[8] Regarding the apparent similarities between primate vocalizations and children's speech, see D. L. Cheney and R. M. Seyfarth, *How Monkeys See the World* (Chicago: University of Chicago Press, 1990), pp. 98–174. Vervet monkeys have different alarm calls for leopards, eagles, and snakes; this shows that vervets perceive at least three different classes of threatening events, which they signal with three different vocalizations. But these are stereotyped and species-universal signals, not independently rehearsable, or refinable; moreover, they are not used representationally, to actively model reality. Thus such calls have none of the critical linguistic features found in the words of children, which indeed are autoretrievable, rehearsable, refinable, and representational.

[9] See D. Premack, *Gavagai* (Cambridge, Mass.: MIT Press, 1987) for an enlightening discussion of this issue.

of human intelligence? A good evolutionary theory of prelinguistic adaptation should try to account for as many of these skills as possible, while providing the cognitive grounds for language.

This leads to my key proposal: the first breakthrough in our cognitive evolution was a radical improvement in human motor control that fortuitously provided a new means of representing reality. *Homo erectus*'s great gift to humanity was "mimetic skill," a revolutionary improvement in voluntary motor control, leading to our uniquely human talent for using the whole body as a subtle communication device. This body-skill might be called pure mimesis, or a talent for action-metaphor. This talent, without language, could have supported a culture that was much more powerful, in terms of its toolmaking abilities, refinements of skill, and flexible social organization than any known ape culture.

For a variety of reasons, mimetic skill logically precedes language, and remains fundamentally independent of truly linguistic modes of representation. It is *the* basic human thought-skill, without which there would not have been the evolutionary opportunity, much later, to evolve language as we know it. Pure mimesis is an intermediate layer of knowledge and culture, and it is the first evolutionary link between the presymbolic knowledge systems of animals and the symbolic systems of modern humans. Basically, mimesis is based in a memory system that can *rehearse and refine movement voluntarily and systematically*, in terms of a coherent perceptual model of the body in the surrounding environment. It is based on an abstract "model of models"[10] that allows any action of the body to be stopped, replayed, and edited under conscious control. This is inherently an autocuing route, since the product of the model is an *implementable self-image*; and although the precise physiological mechanism of this system is not known,

[10] More precisely, an abstract motor model of a very abstract perceptual model.

its functional retrieval path employs kinematic imagery. The principle of retrievability was thus first established at the top end of the motor system; and retrievable body-memories were the first true representations.

Mimesis is by definition a "supramodal" skill. Supramodal means, literally, that mimesis is "above modality," or unrestricted with regard to employing any muscle group in the body toward its ends. A mimetic scenario can be acted out with eyes, hands, feet, posture, locomotion, facial expression, voice,[11] or any other motor modality, or combination of modalities, one can imagine.[12] This is especially evident in the uniquely human behavior pattern known as "rhythm." Rhythm is the motor translation of an abstract sound pattern, the conversion of sound into motion. Rhythm is truly supramodal: rhythmic revelers at a rock concert use every muscle in their bodies to convert an abstract sound pattern into movement. Animals don't have anything like rhythm; apes don't snap their fingers (or anything else) to a beat, and they certainly don't reenact previously heard rhythms on their own, as they move through the forest. They lack the required supramodal mimetic capacity in their brains.

The more complex human motor skills necessitate more than a capacity for supramodal rehearsal: they also require a capacity for purposive sequencing of large-scale patterns of action over longer periods of time, such as those used in advanced toolmaking. This assumes a larger "self-modeling" capacity whereby a series of actions can be imagined and then altered or resequenced.[13] This kind of extended kinematic

[11] At this level, voice control would not yet include phonetic vocalization, and would have been restricted to what we call "prosodic" voice modulation.

[12] This skill is limited, of course, to the skeletomuscular system; smooth-muscle control remains largely reflexive and "involuntary" even in humans. It is fair to say that mimetic skill was built on existing primate skeletomuscular modalities—in other words, all of the modalities subject to the laws of operant conditioning.

[13] A closely related idea was proposed in M. Corballis, *The Lopsided Ape: Evolution of the Generative Mind* (Cambridge, U.K.: Cambridge University Press, 1991). Corballis

imagination is still the basis of human nonverbal imagination, and is essential to the training of those who work with the body, such as actors and gymnasts. Although it is sometimes thought of as primarily visual, nonverbal imagination is more a body-based skill that seems to capture visual images in its wake. It is not an accident that the ancient mnemonic "method of loci"[14] favored by the Greeks and later European cultures did not rely on static visual imagery; rather, it depended on the memorizer generating an image of self-motion inside an imaginary visual space: the kinematic image was thus made the engine of visual recall.

The universality of these uniquely human body skills is still evident in children of all cultures, who routinely practice and refine their motor skills without training or conditioning; images of boys bouncing a ball off the wall over and over, or girls skipping rope endlessly, come to mind. Our close relatives the great apes just don't do that; they do not rehearse and refine action. Nor do other advanced mammals, like dogs and cats. If a dog tries to cross a stream, misses, and slips down the embankment, it doesn't spend the afternoon improving its jumping skill. Baboons throw projectiles in a fight, but they don't systematically practice and improve their throwing skills.

An advance in human motor representation of this magnitude would have automatically had ramifications in the area of expressive capacity. Actions and events could be represented and reenacted independently of the environment; this seems to have showed immediately in much-improved toolmaking and tool use, and in constructional and other instrumental skills. But, as in many evolutionary adaptations, mimetic skill would have had unforeseen consequences: now hominids had a means of re-presenting reality to themselves and others by the use of voluntary action. This means that

argues that an early human motor skill called "generative praxis" enabled *Homo erectus* to achieve a more advanced culture without language. However, generative praxis has quite different properties from mimetic skill.

[14] See J. D. Spence, *The Memory Palace of Matteo Ricci* (New York: Penguin, 1984).

hominids could do much more than rehearse and refine existing movement patterns; they could also imagine and invent completely new patterns of movement, for example, as human gymnasts, dancers, actors, and divers still do. And they could reenact events and scenarios, creating a sort of gestural prototheater of everyday life. The body became a tool for expression, as in acting or mime; it was just a matter of discovering the social utility of this possibility.

The expressive and social aspect of human mimetic skill may be called "pure mimesis." I believe that for a long time (over a million years) early hominids must have subsisted on a "mimetic" culture. Such a culture was based on improved voluntary motor skill, extensive use of imitation for pedagogy, and a much more sophisticated range of voluntary facial and vocal expressions, along with public action-metaphor, the basis of most custom and ritual. Could such a language-less culture have carried *Homo erectus* to the heights he achieved? I have had some serious dissent on this point—for instance, Dennett[15] disagrees with me on the question of whether *Homo erectus* could have supported complex skills like continuous fire-tending without language—but I find strong support for the power and autonomy of nonverbal mind in the study of modern humans.

One line of evidence is the enduring cultural autonomy of mimesis. There are whole areas of modern human culture that continue to function magnificently with a minimal use of language. These include the practice and teaching of many simple crafts and occupations, informal games, especially children's games; many aspects of custom, social ritual, and complex interactive expressive scenarios such as those documented by Eibl-Eibesfeldt[16]; athletic skill and related areas like

[15] D. Dennett, "The Role of Language in Intelligence," Darwin Lecture, Cambridge University, March 6, 1992. (Publication ccs-92-3 of the Center for Cognitive Studies, Tufts University.)

[16] I. Eibl-Eibesfeldt, *Human Ethology* (New York: Aldine de Gruyter, 1989). See also M. Argyle, *Bodily Communication* (London: Methuen, 1975).

fighting skills; and many group expressive customs—for instance, the systematic use of group laughter as a means of ostracism or punishment, and culture-specific customs for indicating deference, affection, manliness, femininity, tolerating pain, celebrating victory, maintaining group solidarity, and so on. These aspects of culture do not depend on language skill, either in their original invention or in their transmission from one generation to the next.

Another line of evidence is neurobiological; these areas of skill are typically resilient in certain cases of global aphasia. This is especially clear in temporary aphasias caused by some types of epilepsy, where patients may lose all use of language (including inner speech) for a few hours, but remain conscious and able to function fairly well on a nonsymbolic level; they can still find their way around in a purposive manner, operate a relatively complex device like a radio or elevator, and manage mimetic social communication (for instance, they know when they are having a seizure and can communicate this to others by gesture).[17] One well-documented patient retained perfect episodic recall for most of his seizure periods; he could describe what went on during the seizure, including exactly who entered and left the room, and what he did with his time. This implies that neither his formation of retrievable memory representations, nor his active recall of them later, could have depended on having a functioning language system. Nor could his perfectly functional mimetic skills have depended on language. The conclusion must be that his mimetic skills come from an autonomous level of representation in the brain, unaffected by his temporary but complete loss of language.

Further evidence for the independence of pure mimesis

[17] The original history on which I based this conclusion, known as the case of "Brother John," was published by A. R. Lecours and Y. Joanette in *Brain and Language* 10 (1980): 1–23. Several other case histories have recently come to my attention, including that of a well-known American cognitive scientist whose introspections were particularly useful because of his cognitive training; he confirmed all of the major points made about Brother John.

comes from the documented lives of illiterate deaf-mutes from the eighteenth and nineteenth centuries, before the diffusion of formal sign languages. Without any effective training regimens to help them communicate, such individuals had to survive without any of the lexical, syntactic, or morphological features of language. They couldn't hear, and thus couldn't have had a sound-based lexicon of words; they obviously lacked an oral lexicon; they couldn't read or write and thus lacked a visually based lexicon. In the absence of a deaf community with a formal sign language, there was no signing lexicon either. Thus none of the lexical components of language were available, and this would have eliminated the possibility of constructing anything we might recognize as true linguistic representations. Yet they often lived remarkable lives,[18] and by recorded accounts were quite sophisticated in their use of pure mimesis, both in constructional skill and in communicative and metaphoric gesture.

In conclusion, mimetic representation is an autonomous, uniquely human level of mind that still supports the nonlinguistic cognitive infrastructure of human society. It allowed humans to break the primate mold and construct retrievable memory representations for the first time. It also led to a very slow-moving process of cultural innovation and change, culminating in the distinctively human cultures of the late period of *Homo erectus*, and setting the stage for a second drastic innovation that would create a much more powerful representational device.

The Second Breakthrough: Lexical Invention. Archaeological markers indicate that a long transition period, from 500,000 to 100,000 years before present, preceded the appearance of modern *Homo sapiens*. During that time a variety of archaic sapient types emerged, and the modern vocal tract began to

[18] See Harlan Lane, *When the Mind Hears* (New York: Random House, 1984).

take form.[19] This is the period when language is most likely to have evolved, culminating in the high-speed speech skills of modern *Homo sapiens*. Language involves a very different type of cognitive operation from the holistic motor strategy underlying pure mimesis. It depends primarily upon a capacity for inventing and retrieving thousands of lexical items—words—along with the rules that govern their use, and constructing narrative commentaries out of these lexical items. Words were the first true symbols, and language in this sense is the signature of our modern human species, *Homo sapiens*. The original modality of language was undoubtedly speech.

There is nothing quite like speech in the world of animal communication; other species, all the way up the evolutionary scale from bees to apes, seem to be limited to a few dozen expressions at most, while humans carry around tens of thousands, and in the case of some multilinguals, hundreds of thousands of words. The cognitive core of this remarkable adaptation is lexical invention; words are portable, highly efficient symbols, and once humans could symbolize aspects of the environment, however haltingly at first, the selection advantages would have been tremendous, especially in an advanced mimetic culture where the complexity of social life was increasing.

Evolutionary pressures favoring a very powerful representational device like speech would have been much greater once a mimetic communicative environment reached a critical degree of complexity; mimesis is inherently an ambiguous way of representing reality, and words are an effective means of disambiguating mimetic messages. Modern children still acquire speech in this way, with most of their early utterances embedded in mimetic exchanges, usually with the mother. The latter include pointing, tugging, prosodic voice sounds, eye contact, nonlinguistic sounds and gestures, and mimetic

[19] P. Lieberman, *The Biology and Evolution of Language* (Cambridge, Mass.: Harvard University Press, 1984).

whole-body movement. Even when the young child is "talking" to itself, it is usually in a mimetic context.

Lexical invention is a constant process of labeling, defining, and differentiating aspects of the perceived world (including the products of speech itself): humans are constantly inventing new lexical items or acquiring them from others, and oral languages are seldom static for very long, revealing a continuing tension between lexical inventions and their significations, as if there was a natural tendency for the system to keep differentiating and defining reality. The earliest human languages would have gone through a "protolinguistic" phase, and the form of that protolanguage might well have resembled Bickerton's protolinguistic, grammarless speech acts, but it seems to have arrived much later than he suggested, originating during the past 300,000 years, rather than a million years earlier, and *only after the principle of voluntary self-cued memory retrieval had been firmly established in the human brain.*[20] Grammatical invention appears to be just another product of a general capacity for lexical invention[21] but there is some neurophysiological evidence supporting Bickerton's view that grammatical invention was the "second phase" of language emergence, requiring a different brain adaptation from lexical invention.[22] Like mimesis, language is at core a thought-skill, but rather than using the holistic, quasi-

[20] Viewed as a motor production system, the emergence of speech required the preexistence of vocomimetic skill, in that the very brief vocal acts that are used to assemble words—their phonology—are really "articulatory gestures" with many of the same properties as other voluntary, refinable mimetic skills. See C. P. Brownman and L. Goldstein, "Articulatory Gestures as Phonological Units," *Phonology* 6 (1989): 201–251.

[21] This is compatible with the theoretical position expressed by E. Bates and B. MacWhinney, "Competition, Variation and Language Learning," *Mechanisms of Language Acquisition*, B. MacWhinney, ed. (Hillsdale, N.J.: Erlbaum, 1987).

[22] This evidence shows that grammar-related words (especially the so-called function words) might have a somewhat separate physical basis in the brain from words that are unrelated to grammar (like adjectives and nouns); see H. Neville, "Fractionating Language: Different Neural Subsystems with Different Sensitive Periods," *Cerebral Cortex* 2 (1992): 244–258.

perceptual strategy of mimetic motor skill, it employs true symbols and constructs narrative descriptions of reality.

Spoken language provided humans with a second form of retrievable knowledge and a second-order "model of models" with a much more powerful way to format their knowledge. The natural product of language is narrative thought, that is, storytelling. Storytelling had a forerunner in mimetic event-reenactment, but it is very different in the means by which it achieves its goal, and much more flexible in what it can express. Mimetic reenactment is bound to imagery of the original event being depicted; when actors communicate grief or heroism mimetically, they are limited in the number of ways it can be done, because the representation is tied to the audience's original perception of grief or heroism. The quintessential narrative act—verbally labeling the agents, actions, and their relationships—lifts the observer outside of space and time, allowing the component parts of the story to be examined, reassembled, and shared much more freely.

The incessant speed of oral linguistic invention can be seen in evidence that all known Indo-European languages have differentiated in the brief period since about 4000 B.C.,[23] and that the even more diverse families of native North and South American languages differentiated from three root Asiatic languages during the past 15,000 years.[24] Spoken language altered human culture not merely in the number and complexity of available words and grammars, but in the shared products of oral cultures. The collective use of narrative thought led inevitably to standardized descriptions: shared, agreed-upon versions of past events. These "official" versions of events formed the basis of myth and religion, which were the direct products of evolving

[23] See C. Renfrew, "The Origins of Indo-European Languages," *Scientific American*, October 1989; and T. V. Gamkrelidze and V. V. Ivanov, "The Early History of Indo-European Languages," *Scientific American*, March 1990.

[24] J. H. Greenberg and M. Ruhlen, "Linguistic Origins of Native Americans," *Scientific American*, November 1992.

linguistic skill. It is a telling point that mythic invention seems to have preceded any further advances in human toolmaking; even the most technologically primitive cultures ever documented have had fully developed oral languages and mythic systems. However, the new oral cultures did not abandon mimetic representation; to the contrary, they encompassed the more concrete, pragmatic culture of mimesis, which continued to function much as it had in the past, in its own traditional cultural arenas. Mimetic skill still provides the cognitive basis for human social institutions like craft, athletics, dance, and the complex nonverbal expressive dimensions captured and cultivated in ritual, acting, and theater; and language provides the narrative framework that ultimately governs those institutions.

Myth and narrative thought in general might be called the "governing" level of thought in oral cultures. All humans grow up within a mythic system, whether they know it or not; myths form the cultural glue that holds societies together. Myths and stories contain and supersede the prototypes and mimetic stereotypes of social roles, social structure, and custom. They rely heavily on allegory and metaphor, and they lack precision, but they remain the most universal form of human integrative thought, and one of the most potent and personally meaningful ways of representing reality.

In modern humans, language and mimetic skill usually work closely together in the expression of ideas,[25] but also can be used independently of one another, to create simultaneously contrasting messages. Such contrasts are a common device in many areas of culture, but especially in cinema, theater, comedy, and opera. For instance, Verdi's *Rigoletto* employs mimetic-linguistic counterpoint very effectively: at several points in the opera the libretto is contradicted by an unfolding mimetic scenario; and the tension produced by driving these two contrasting modes of representation in opposite directions

[25] See D. McNeill, *Hand and Mind: What Gestures Reveal About Thought* (Chicago: University of Chicago Press, 1991).

is a very powerful dramatic device. However, from a cognitive standpoint the important observation is that the audience seems to have no difficulty "parallel processing" the contrasting mimetic and linguistic messages. This suggests that these two separate representational realms are sufficiently independent in the brain that they can operate concurrently, without interfering with one another.

The Third Breakthrough: Externalization of Memory

The two evolutionary steps described above form the innate structural foundations of human thought, our gene-based cognitive inheritance. But cognitive evolution did not stop when we reached our modern form somewhere between 50,000 and 100,000 years ago; a third major cognitive breakthrough has to be posited to account for the astonishing changes that have taken place more recently. These changes revolve around one central historical trend that has dominated the history of the past 20,000 years: the externalization of memory.

Early humans, like their predecessors, depended on their natural or biological memory capacities. Thus, even though language and mimetic expression allowed humans to accumulate a considerable degree of collective knowledge shared in culture, the actual physical storage of that knowledge depended ultimately on the internal memory capacities of the individual members of a society. Thought was carried out entirely inside the head; whatever was heard or seen had to be remembered and rehearsed orally or visualized in imagination.

The advantages of external memory storage may now be obvious to us, but the invention of the many external memory devices we now enjoy has taken at least 20,000 years, and the full social realization of the power of external symbols is very recent. The common keyword for the most recent phase of that transformation is "literacy," but this term needs broadening to include much more than conventional literacy, which in Western culture often means simply the ability to read and write alphabetic symbols. A more adequate description of human symbolic literacy would encompass all of the new skills needed to use every kind of permanent external symbol, from the pictograms and line drawings of the Upper Paleolithic, to the astrolabes and alchemical diagrams of the medieval era, to the digital information codes used in modern electronic communications.

There has been no time for a genetic adaptation for external symbol use. We are equipped with basically the same brain we had about 50,000 years ago, and it might be argued that the shift to external memory was purely cultural, and therefore not as fundamental a change as the two previous ones. However, using the same criteria used to evaluate earlier major cognitive changes, recent changes constitute strong evidence for a third major breakthrough in our cognitive evolution. Like the two previous ones, the physical medium as well as the functional architecture of human memory have changed, and new kinds of representations have become possible.

External symbols have radically changed the medium of storage. In this, they constitute a real hardware change in memory, albeit a change in technological rather than in biological hardware. This change was not trivial, since the storage properties of external media are very different from those inside the head. Whereas biological memory records are in an impermanent, fixed medium with a constrained format, external records are usually in enduring, refinable, and reformatable media. These properties allow humans to construct completely new types of memory records.[26] They also allow us to greatly expand the amount of knowledge stored in memory. The size of single entries in biological

[26] I have suggested these might be called "exograms" to complement the corresponding term "engrams" that Karl Lashley coined to label biological memory records. The properties of exograms and external memory networks are discussed at more length in my book *Origins of the Modern Mind* (Cambridge, Mass.: Harvard University Press, 1991), pp. 308–333.

memory is very limited, and the total number of entries is also limited; these constraints do not apply to external storage.

External storage has also introduced new ways of retrieving and organizing information; in fact the retrieval workhorses of biological memory—similarity, and temporal and spatial contiguity—are not particularly important in external memory retrieval. And the addition of so many external devices has actually changed our memory architecture—that is, the storage and processing options in the system, and their configuration—allowing us to move freely through an external information space that is spatialized, or virtually frozen in time. Because of their stable display properties, external memory devices have allowed us to harness the power of our perceptual systems, especially vision, for reflective thinking; and they have literally changed which part of the brain we use to do much of our thinking. They have also increased our options for interrelating various kinds of images and information, and for doing mental work in groups.

Moreover, there is a neuropsychological dimension to all of this. This can be seen more clearly in historical perspective. Once, there were virtually no external symbols; then few; and then humans were rather rapidly surrounded by the thousands of symbolic codes and conventions that mark modern society. This has amounted to a virtual invasion of the brain by culturally imposed programming, mostly in the form of institutionalized education. An adequate description of this tremendous transition to symbolic literacy should include all of the internal programming needed to manage this massive culturally driven load on the brain.

A partial list of devices mastered by humans along the way to full symbolic literacy includes (in rough historical order) iconography, maps, emblems, totems, pictorial representations, pictographs, sequence-markers like knotted cords or prayer beads, various types of tokens, currencies, property markers, ideographic writing systems, counting systems, mathematical notations, schematic and geometric diagrams, lists, syllabaries and alphabets, scrolls, books, archival records of various sorts, military plans, organizational charts, environmental signs of various kinds, graphic images, scientific manuals, graphs, analog instruments, specialized technical languages, computing languages, and a variety of modern multimedia storage devices that employ virtually all of the above. A modern professional uses thousands of such devices in one day, moving from one externally determined situation to another, flipping from book to report to map to newspaper to computer to the dictionaries and manuals that point their way through the signaled universe, guided all the while by legions of environmental directives and signals. Even the personal memory system has been programmed to the hilt with photographs, memoranda, TV images, and other kinds of stored knowledge.

The load placed on high-level visual skills is especially striking. Most external storage devices are visual, and require a second level of visual processing that is missing in completely nonliterate minds. In addition to the "natural" visual encoding of things like environmental objects and complex social events, a level of symbolic encoding is required of external memory devices; in effect they must be processed in two ways, both as natural objects or events and as symbolic artifacts. Thousands of internal visual-symbolic codes provide the linkages needed to provide the correct semantic reference, or meaning, for a given symbolic device. The arbitrariness of these linkages is evident on examining the hundreds of different varieties of writing that humans have invented: the same "meaning" can be addressed in a virtually infinite number of ways.

Once the required codes are in place in the brain, and assuming that the semantic memory system has a sufficient base of knowledge to work with, a successful external memory device will *reproduce an intended mind-state* in the reader or viewer. To the expert reader, the encoding strategies are so deeply established that the medium itself is invisible; ideas literally pop out of the page, and the processing of the message is unconscious. While processing a major symbolic artifact—a novel, for instance—a particular set of abstract representations will be set up in the reader's mind; and this temporary mind-state is highly dependent on the external device. Once the artifact is taken away, very little remains; put down a long novel and the temporary richness of the story will very rapidly dim, usually leaving only a very general impression of the story and its characters. Pick it up again, and within minutes the "world" created by the author will reproduce itself in the reader's mind.

The literate mind has thus become externally programmable, which is both an advantage and a danger. The advantage lies in the creative possibilities of symbols; societies can support much greater complexity, science and technology can advance, scholarship becomes possible, and artists and writers become cognitive engineers, leading their audiences through tangled symbolically driven nets of ideas to end-states that are not otherwise conceivable. The disadvantage is found in potential threats to individual integrity; free access to external memory tends to pull apart the unity of mind, fragmenting experience, undermining the simpler mythic thought structures humans have grown rather attached to, and exposing them to a bewildering variety of very powerfully packaged messages.

The extent of the load imposed on the brain by full literacy has never really been appraised, but an estimate might be made simply by listing the number of different exemplars of each type of symbol a highly literate mind should recognize. In Western society a great deal of emphasis is placed on the ability to translate print into sound, and vice versa, using phonetic rules. Estimates based on written vocabulary samples suggest that college-educated readers of English probably have about 75,000 words in their reading lexicons,[27] and multilingual people may have many more than this, particularly if they have

[27] R. C. Oldfield, "Individual Vocabulary and Semantic Currency: A Preliminary Study," *British Journal of Social and Clinical Psychology* 2 (1963): 122–130.

to read in several very different scripts. One former director of language services for the United Nations claimed nineteen fluent languages and another twelve that were "a bit rusty." These included some very diverse language groups: some Chinese and Native American languages, as well as some of the Germanic, Latin, Arabic, and Celtic languages. This would suggest that he recognized at least half a million lexical entries, while keeping separate the enormously complex symbolic conventions of several major writing systems and some of their local variants.

This is an extreme case, but we should not underestimate the coding demands placed on the minds of our most highly trained professionals; and these demands extend well beyond conventional concepts of alphabetic, or phonetic, literacy. They include pictorial symbols which cannot be understood unless the viewer is familiar with their meaning in a particular society—thus to be an art historian one needs a vast pictorial "lexicon," and so do afficionados of pop culture or of cinema. There are also extensive ideographic lexicons in most literate brains, even in societies that use alphabetic print; most adults can recognize thousands of ideograms, including those used in maps, road signs, advertisements, and scientific data. Much more elaborate ideographic lexicons are required in societies that use writing systems like Chinese, which is essentially ideographic and therefore may be "read" in dozens of different spoken languages. Numbers and mathematical equations share this property: the same balance sheet or statistical table can be read in New York, Tehran, or Tokyo, because the tables represent ideas directly rather than by mapping onto sound. A degree of ideographic competence was always a major part of literacy.

The more complex forms of symbol use require the combination of all of these forms of visual representations—pictorial, ideographic, and phonetic—into large-scale external artifacts like architectural proposals, engineering plans, government reports, scientific treatises, cinematic scripts, or works of art. The high-level or "metalinguistic" skills needed to do this kind of mental work are difficult to acquire, and far from being universal to all humans.

These brave new capacities were not acquired without tradeoffs. There is only so much brain-matter (or more properly, mind-matter). The physiology of brain plasticity, or malleability, suggests that when we increase the demands on one area of the brain it expands its territory more or less in proportion to the imposed load.[28] Accordingly, cerebral capacity is used up and is no longer available for something else; there are no freebees in the neurobiological economy, any more than elsewhere. There is some evidence that with literacy we sacrificed a degree of visual imagination; and that we are losing our capacity for rote verbal skills, like mental arithmetic and memorization.[29] The nature of these tradeoffs should be explored further, because it is a certainty that symbolic literacy isn't easily or naturally acquired in development, the way mimesis and speech are. Rather, literacy is most unnatural, and requires a wrenching redeployment of cerebral resources.

The most complex artifacts of external memory have always been accessible only to a fraction of the population, but for a long time they have been at the center of power. Even the earliest government bureaucracies needed some form of writing for trade, census-taking, and taxation, and this was where formal education really began; roughly 15 percent of all the ancient cuneiform tablets from Uruk and Babylon were lexical lists for training scribes. The scale of this educational effort was very limited; scribes alone knew the secret codes for using symbols. But it was this small class of people that developed and standardized, over thousands of years, most of

[28] See, for example, M. M. Merzenich et al., "Variability in Hand Surface Representations in Areas 3b and 1 in Adult Owl and Squirrel Monkeys," *Journal of Comparative Neurology* 258 (1987): 281–296.

[29] There is an especially long line of studies on the apparent loss of imagery with age and schooling; see, for instance, A. Richardson, *Mental Imagery* (New York: Springer, 1969).

the systems of external symbolic representation that we know as writing. Literacy governed many ancient societies, but this process was largely indirect, and the vast majority of people, even in Greece and Rome, remained illiterate, carrying out their written transactions through professional scribes.[30] While they may not have been literate, however, their lives, and many of their ideas, including their global spatiotemporal models of the world, were the products of symbolically literate minds. As work requiring literacy skills has become more common, the power of the highly literate elite has continued to expand rather than shrink. And as technology accelerates and deepens the process of symbolic governance, the consequences of exclusion from access to the products of symbolic memory become more and more serious. Claims of the demise of literacy are exaggerated; if anything, society now depends on symbolic literacy skills more than it ever has.

Changes in the Role of Biological Memory

External memory technologies have radically changed the role, and organization, of biological memory. In all animals, including the human animal, biological memory is usually divided into a "long-term" store that contains a lifetime of acquired knowledge, and a "working memory" that serves as a temporary holding place for the immediate demands of the moment. Working memory is in the background of consciousness, and central to what we experience as conscious thought; in fact many cognitive theorists have argued that working memory is the central locus of thought. Alan Baddeley of Oxford University has reviewed evidence on this question, and concluded that the human "working memory" has a tripartite structure, within which most of our conscious mental work

[30] W. V. Harris, *Ancient Literacy* (Cambridge, Mass.: Harvard University Press, 1989).

takes place.[31] This tripartite structure involves a central executive memory and two slave systems that he calls the visual sketchpad and the articulatory loop. Basically, when we think, we either imagine (via the sketchpad) or mutter to ourselves (the articulatory loop); inner speech is simply a covert version of the articulatory loop. The central executive deals with more abstract ideas rather than sensory impressions: thus in a long conversation, the central executive keeps track of what was said, by whom, and in what context, not in literal detail, but on the level of meaning. The same would apply to one's own thoughts; the central executive maintains the general semantic context.

In preliterate cultures this natural working memory structure was all we had to work with, and it is a very fragile, limited structure.[32] Hence the tendency of preliterate oral cultures to lean heavily on literal verbal memory, formulaic recital, rigid group ritual, and visual imagination. This is not to underestimate the considerable intellectual achievements of oral culture—for instance, rainforest pharmacopoeias and Micronesian celestial navigation—but there are severe limits to how far such cultures have historically been able to develop.

Permanent external symbols changed this cognitive arrangement. With external symbols we have an *external* memory field, in the form of an immediate display of selected artifacts, that often serves as our real working memory. The thinker holds the displayed item in the external field, and plays with that item in iterative loops, improving or extending the memory representation in the external memory field. The primary

[31] A. Baddeley, *Working Memory* (Oxford: Clarendon Press, 1986).

[32] Stripped of all our external symbols, our working memory system is vulnerable to interference and cannot hold more than a few items at a time; moreover, without constant rehearsal it can hold these few items for only a few seconds. It has to process inputs serially, that is, one item at a time, and it must keep moving. Ideas cannot be frozen in time as they are in external symbolic devices, and therefore there is too short a dwell-time on any specific item to allow any prolonged reflection. See Donald, *Origins of the Modern Mind*, pp. 325–333.

display field may be centered on a computer screen, but it usually includes many other items as well. This active interaction with a display field has become an essential tool for modern styles of thinking, and amounts to a major change in the schematic or architectural diagram of human working memory.

Most of our "long-term storage" is now external rather than biological. The omnipresence of external memory has changed the way we perform mental work in collectivities. There has been a recent burst of research on the topic of "distributed cognition," that is, on how cognition is increasingly becoming a shared activity, in working groups linked together by electronic memory technology. For example, Hutchins[33] studied the cockpits of airplanes as integrated information systems in which no single element, neither humans nor the instruments, have all of the information; they work together as a unit, and it is more than a truism to say that such a system knows more than any of its components, as least with reference to its operational function. Anthropological studies of cooperative computer-coordinated work tell a similar story; teams of executives or information specialists play out collective scenarios in which the behavior of individuals can be understood only as part of a larger interactive system. Managers are working on ways to control and predict the behavior of such collectivities.[34] Thus the new anthropology is not being carried out in some isolated exotic locale; it has moved to the boardrooms and workplaces of the emerging global metropolis, where the really esoteric new forms of cognition are emerging.

[33] E. Hutchins, "The Technology of Team Navigation," *Intellectual Teamwork: Social and Technological Foundations of Cooperative Work*, J. Galegher, R. E. Kraut, and C. Egido, eds. (Hillsdale, N.J.: Lawrence Erlbaum Associates, 1985), pp. 191–220.

[34] See J. R. Olson and G. M. Olson, "User-Centered Design of Collaboration Technology," *Journal of Organizational Computing* 1 (1991): 61–83; also L. Suchman, *Plans and Situated Actions: The Problem of Human-Machine Communication* (New York: Cambridge University Press, 1987).

And now that we can store, display, and refine the forms of ideas externally, and no longer depend so heavily on biological working memory (either for storage, or for retrieval and display), it will be important to determine more precisely what is happening to biological working memory as the use of external memory expands. Given our symbiosis with external devices, the content of human awareness is more and more determined by external symbols, or directly controlled from an external memory field, while internal working memory falls into a subsidiary managerial role. Wherever new memory technology may be leading us, it is difficult to imagine a more fundamental change in our cognitive situation.

In conclusion, humans alone have gained voluntary access to their memory systems, and humans alone can craft memory—that is, create representations. There are three broad realms of representation, each fundamentally independent of the other. The first (pure mimesis) is based on holistic motor models of event-perceptions. The second (language) invents true symbols and constructs narrative descriptions. The third (external memory) contains symbolic devices that can set up specific mind-states in a "literate" recipient. The modern mind is in perpetual tension between these three great realms—the mimetic theater of action; the world of verbal description; and the external symbolic storage system. The third realm is expanding exponentially, and this is a challenge to traditional culture, which is based on the first two. It is also a challenge to the human nervous system, which has limited resources and is increasingly pressured to trade off traditional competencies against those imposed by the demands of external memory.