THE HUMAN ADAPTATION FOR CULTURE

Michael Tomasello
Max Planck Institute for Evolutionary Anthropology, Inselstrasse 22, D-04103 Leipzig, Germany; e-mail: tomas@eva.mpg.de

Key Words: culture, cognition, ontogeny, primates, language

Abstract Human beings are biologically adapted for culture in ways that other primates are not, as evidenced most clearly by the fact that only human cultural traditions accumulate modifications over historical time (the ratchet effect). The key adaptation is one that enables individuals to understand other individuals as intentional agents like the self. This species-unique form of social cognition emerges in human ontogeny at approximately 1 year of age, as infants begin to engage with other persons in various kinds of joint attentional activities involving gaze following, social referencing, and gestural communication. Young children’s joint attentional skills then engender some uniquely powerful forms of cultural learning, enabling the acquisition of language, discourse skills, tool-use practices, and other conventional activities. These novel forms of cultural learning allow human beings to, in effect, pool their cognitive resources both contemporaneously and over historical time in ways that are unique in the animal kingdom.

CONTENTS
Introduction ........................................................................................................ 510
Primate and Human Cognition ........................................................................... 511
The Ontogeny of Human Cultural Learning .................................................. 513
Joint Attention .................................................................................................... 513
Imitative Learning ............................................................................................. 514
Linguistic Symbols and Cognitive Representation ........................................ 516
Nonhuman Primate Culture ............................................................................. 518
Japanese Macaque Potato Washing ............................................................... 519
Chimpanzee Tool Use ..................................................................................... 519
Chimpanzee Gestural Communication ............................................................ 521
Nonhuman Primate Social Learning and Social Cognition .......................... 523
Encultuated Apes ............................................................................................ 524
Conclusions ........................................................................................................ 526
INTRODUCTION

All animal species have unique characteristics and human beings are no exception. Perhaps most important, human beings have some unique cognitive skills. The precise nature of these skills is unknown, but they must be such that they enable a number of species-unique and easily observable behavioral practices, including the following: (a) the creation and use of conventional symbols, including linguistic symbols and their derivatives, such as written language and mathematical symbols and notations; (b) the creation and use of complex tools and other instrumental technologies; and (c) the creation of and participation in complex social organizations and institutions. It is difficult to imagine a more fundamental anthropological question than that of where these complex and species-unique behavioral practices, and the cognitive skills that underlie them, came from.

Recent research on human evolution has provided some important facts that may help us to answer this most basic of questions. First, human beings shared a common ancestor with their nearest primate relatives, Pan troglodytes and Pan paniscus, a mere 6 million years ago—about the same time there existed a common ancestor for horses and zebras, lions and tigers, and rats and mice (King & Wilson 1975). Second, for almost two thirds of this 6 million years, the human lineage consisted of one or more species of Australopithicine, which in most recent research are characterized as ape-like in both brain size and behavior (Klein 1989). Third, although controversial, many anthropologists now believe that in the 2 million years of the existence of genus Homo, it has only been during the past several hundred thousand years, with the rise of something like modern humans, that the unique aspects of human cognition have come into full bloom (Stringer & McKie 1996). What these new facts and interpretations establish is the rapidity with which the species-unique aspects of human cognition must have arisen: within the past 6 million years for certain, within the past 2 million years in all likelihood, and within the past half million years according to some respectable theories. The main point is that under none of these scenarios—especially the last—has there been sufficient time for a large number of major cognitive adaptations (contra most of so-called evolutionary psychology) (Tooby & Cosmides 1989, Pinker 1997). If we are searching for the origins of uniquely human cognition, therefore, our search must be for some small difference that made a big difference—some adaptation, or small set of adaptations, that changed the process of primate cognitive evolution in fundamental ways.

In my view there is only one candidate for this small difference that made a big difference and that is human culture. Other primates and mammals are certainly social, and some may even have social organizations for which it is useful to apply the term culture (McGrew 1998). But human social organization is something else again, and this organization was, in my view, an integral part of the process by which human cognition came to have many of its most distinctive characteristics. That is, although the cognition of many mammalian and primate species is influenced in important ways by their social environments, human cognition, at least in its species-unique aspects, is actually socially constituted. In this
paper, I attempt to explicate this proposition more fully and to explore some of its most important anthropological implications by systematically comparing the social learning, social cognition, and cultural organization of human beings and their nearest primate relatives.

PRIMATE AND HUMAN COGNITION

Human cognition is a species, in the literal meaning of the word, of primate cognition. Tomasello & Call (1997) reviewed all of the most important studies of primate cognition over the past century and established that a vast array of cognitive skills are common to all primates, including humans. Thus, in their cognition of the physical world, all primate species remember “what” is “where” in their local environments, take novel detours and shortcuts in navigating through space (cognitive mapping), predict where food will be located in the future based on a number of current cues, follow the visible and invisible displacements of objects (Piaget’s object permanence), categorize objects on the basis of perceptual similarities, understand relational categories and perform mental rotations of objects in space, match (and perhaps add) small numerosities of objects independent of spatial cues, and use creative strategies and perhaps insight in problem solving (sometimes in tool use). The major conclusion is that all primates live in basically the same sensory-motor world of permanent objects—and categories and relations of permanent objects—arrayed in a representational space, and they all have some insightful problem-solving skills (in some cases involving the making and using of tools) to affect that sensory-motor world.

There are also many similarities in the way all primate species understand their social worlds. Thus, all primate species recognize individuals in their social groups; form direct relationships with other individuals based on such things as kinship, friendship, and dominance rank; understand the third-party social relationships that hold among other members of their groups—again based on such things as kinship, friendship, and dominance rank; predict the behavior of individuals based on a variety of social and physical cues, and sometimes novel insights; use many types of social and communicative strategies to solve social problems and so to out-compete groupmates for valued resources; cooperate with conspecifics in problem-solving tasks and in forming social coalitions and alliances; and engage in various forms of social learning, in which they learn valuable things from conspecifics. The major conclusion again is that all primates live in basically the same type of social world, in which they individually recognize conspecifics and appreciate both the vertical (dominance) and horizontal (affiliative) relationships that hold among group members. They also have the ability to predict the behavior of conspecifics in many situations based on a variety of cues and insights, and in some cases to affect the behavior of groupmates via various social and communicative strategies.

So what makes human cognition different? What enables human beings to create and use language and other symbols, to create and maintain complex instrumental technologies, and to create and maintain complex social organizations and
institutions? The first and most obvious observation is that individual human beings do not do any of these things. These are all collective cognitive products in which human beings have in some way pooled their cognitive resources. If we imagine the forbidden experiment in which a human child grows up on a desert island, miraculously supplied with nutritional and emotional sustenance but in the total absence of contact with other human beings, this child would not invent a language or a complex technology or a complex social institution. Even if there were a group of such abandoned children, it is unlikely that together, in their own lifetime, they could invent anything resembling the range of material and symbolic artifacts that characterizes even the least artifactually complex human society. The reason that no single child or group of children could on their own in their own lifetimes create any version of a modern human culture and its material and symbolic artifacts is that human cultures are historical products built up over many generations. Indeed, the most distinctive characteristic of human cultural evolution as a process is the way that modifications to an artifact or a social practice made by one individual or group of individuals often spread within the group, and then stay in place until some future individual or individuals make further modifications—and these then stay in place until still further modifications are made (Tomasello et al 1993a, Boesch & Tomasello 1998). This process of cumulative cultural evolution works because of a kind of “ratchet effect”: Individual and group inventions are mastered relatively faithfully by conspecifics, including youngsters, which enables them to remain in their new and improved form within the group until something better comes along.

The major part of the ratchet in the cumulative cultural evolution of human societies takes place during childhood. That is, each new generation of children develops in the “ontogenetic niche” characteristic of its culture (including in some cases explicitly pedagogical niches), mastering the artifacts and social practices that exist at that time. It is only because human children are so good at social learning (and in some cases adults are so good at teaching) that an artifact or social practice may conserve its form over many generations of stasis, until eventually a modification that group members find worthwhile is made and the cycle starts anew. For this process to work, therefore, human beings not only need to be inventive, they need to be good at preserving those inventions by imitatively learning, and sometimes explicitly teaching, the inventions of others. This process is more complex than it might seem at first glance, however. Imitative learning does not just mean mimicking the surface structure of a poorly understood behavior, the way a parrot mimics human speech, with no understanding of its communicative significance, it also means reproducing an instrumental act understood intentionally, that is reproducing not just the behavioral means but also the intentional end for which the behavioral means was formulated. This requires some specially adapted skills of social cognition.

The main point is that unlike the young of any other primate species, human children grow up in the midst of the accumulated wisdom of their social group, as embodied in its material artifacts, symbolic artifacts, and conventional social practices, and children are specifically adapted to appropriate this wisdom as
embodied in these forms. Although we have yet to explore the details, these facts provide a sufficient explanation for the existence of many of the most distinctive cognitive products that human beings produce. But there is more. As a result of participating in social and communicative interactions with other persons understood intentionally, human children also come to cognitively represent the world in some uniquely powerful ways. The most important of these involves the use of linguistic symbols that are both intersubjectively understood and perspectival, in the sense that that a single item or situation may be construed linguistically in many different ways. For example, a single event may be seen as X sold Y to Z or Z bought Y from X; or a single object may be an apple, a fruit, or some food. Intersubjective and perspectival cognitive representations are unique in the animal kingdom, and they enable human beings to deal with their worlds in some uniquely flexible and powerful ways.

To appreciate fully (a) the social cognitive skills necessary for children to participate fully in their cultures and (b) the transforming effect of cultural participation on individual cognition, we must follow out key aspects of the human ontogenetic scenario—and then compare it with the basic scenario of nonhuman primate cognitive ontogeny.

THE ONTOGENY OF HUMAN CULTURAL LEARNING

Human children grow up in the ontogenetic niche of their culture, which, in a sense, exists before they are born. But children also need to have some social cognitive skills if they are to exploit the preexisting cultural resources in a species-typical manner. These skills cannot be simply presupposed, as is often the case in cultural psychology (e.g. Rogoff 1990, Shweder 1990). This point is most clearly demonstrated by the unfortunate case of children with autism, the vast majority of whom lack the social cognitive skills necessary to participate fully in or to appropriate the artifacts and social practices characteristic of those around them (Hobson 1993, Baron-Cohen 1993). For typically developing children, the ontogeny of these social cognitive skills begins at the end of the first year of life.

Joint Attention

Six-month-old infants interact dyadically with objects, grasping and manipulating them, and they interact dyadically with other people, expressing emotions back-and-forth in a turn-taking sequence. But at approximately 9–12 months of age, infants begin to engage in interactions that are triadic in the sense that they involve the referential triangle of child, adult, and some outside entity to which they share attention. Thus, infants at this age begin to flexibly and reliably look where adults are looking (gaze following), use adults as social reference points (social referencing), and act on objects in the way adults are acting on them (imitative learning)—in short, to “tune in” to the attention and behavior of adults toward outside entities. At this same age, infants also begin to use communicative gestures to direct adult attention and behavior to outside entities in which they are interested—in short, to get the adult to “tune in” to them. In many cases, several of
these behaviors come together as the infant interacts with an adult in a relatively extended bout of joint engagement with an object (Bakeman & Adamson 1984). Most often the term joint attention has been used to characterize this whole complex of triadic social skills and interactions (Moore & Dunham 1995), and it represents something of a revolution in the way infants relate to their worlds.

Infants begin to engage in joint attentional interactions when they begin to understand other persons as intentional agents (Tomasello 1995). Intentional agents are animate beings with the power to control their spontaneous behavior, but they are more than that. Intentional agents also have goals and make active choices among behavioral means for attaining those goals. It is important to note that intentional agents also make active choices about what they pay attention to in pursuing those goals (for the argument that attention is intentional perception, see Gibson & Rader 1979). All of the specific joint attentional behaviors in which infants follow, direct, or share adult attention and behavior are not separate activities or cognitive domains, they are simply different behavioral manifestations of this same underlying understanding of other persons as intentional agents. Strong support for this view comes from a recent study by Carpenter et al (1998b), who followed a group of infants longitudinally from 9 to 15 months of age and found that for any individual child these skills emerged together as a group, with some predictable orderings among individual skills.

Imitative Learning

The social-cognitive revolution at 1 year of age, sets the stage for infants’ second year of life, in which they begin to imitatively learn the use of all kinds of tools, artifacts, and symbols. For example, in a study by Meltzoff (1988), 14-month-old children observed an adult bend at the waist and touch its head to a panel, thus turning on a light. They followed suit. Infants engaged in this somewhat unusual and awkward behavior, even though it would have been easier and more natural for them simply to push the panel with their hand. One interpretation of this behavior is that infants understood that (a) the adult had the goal of illuminating the light and then chose one means for doing so, from among other possible means, and (b) if they had the same goal, they could choose the same means. Cultural learning of this type thus relies fundamentally on infants’ tendency to identify with adults, and on their ability to distinguish in the actions of others the underlying goal and the different means that might be used to achieve it. This interpretation is supported by the more recent finding of Meltzoff (1995) that 18-month-old children also imitatively learn actions that adults intend to perform, even if they are unsuccessful in doing so. Similarly, Carpenter et al (1998a) found that 16-month-old infants will imitatively learn from a complex behavioral sequence only those behaviors that appear intentional, ignoring those that appear accidental. Young children do not just mimic the limb movements of other persons, they attempt to reproduce other persons’ intended actions in the world.

Although it is not obvious at first glance, something like this same imitative learning process must happen if children are to learn the symbolic conventions of their native language. Although it is often assumed that young children acquire
language as adults stop what they are doing, hold up objects, and name these objects for them, this is empirically not the case. Linguistics lessons such as these are (a) characteristic of only some parents in some cultures and (b) characteristic of no parent in no culture for words other than concrete nouns and some actions, i.e. no one names for children acts of “giving” or prepositional relationships such as “on” or “for.” In general, for the vast majority of words in their language, children must find a way to learn in the ongoing flow of social interaction, sometimes from speech not even addressed to them (Brown 1999). In some recent experiments, something of this process has been captured, as children learned words in situations in which the adult was not specifically intending that they learn a word, the referent was not perceptually available when the word was said, and there were multiple potential referents in the situation that the child had to choose among based on various kinds of adult social-pragmatic cues.

1. In the context of a finding game, an adult announced her intentions to “find the toma” and then searched in a row of buckets all containing novel objects. Sometimes she found it in the first bucket searched, smiling and handing the child the object. Sometimes, however, she had to search longer, rejecting unwanted objects by scowling at them and replacing them in their buckets until she found the one she wanted (again indicated by a smile and the termination of search). Children learned the new word for the object the adult intended to find regardless of whether or how many objects were rejected during the search process (Tomasello & Barton 1994).

2. Also in the context of a finding game, an adult had the child find four different objects in four different hiding places, one of which was a distinctive toy barn. Once the child had learned which objects went in which places, the adult announced her intention to “find the gazzer.” She then went to the toy barn, but it turned out to be “locked.” She thus frowned at the barn and then proceeded to another hiding place, saying “Let’s see what else we can find” (taking out an object with a smile). Later, children demonstrated that they had learned “gazzer” for the object they knew the experimenter wanted in the barn—even though they never saw the object after they heard the new word, and even though the adult had frowned at the barn and smiled at a distractor object (Akhtar & Tomasello 1996, Tomasello et al 1996).

3. An adult announced her intention to “dax Mickey Mouse” and then proceeded to perform one action accidentally and another intentionally (or sometimes in reverse order). Children learned the word for the intentional not the accidental action regardless of which came first in the sequence (Tomasello & Barton 1994).

Tomasello et al (1993a) called this kind of imitative learning cultural learning because the child is not just learning things from other persons, she is also learning things through them—in the sense that she must know something of the adult’s perspective on a situation to learn the active use of this same intentional act. The adult in the above scenarios is not just moving and picking up objects randomly, she is searching for an object and the child must know this in order to
make enough sense of her behavior to connect the new word to the adult’s intended referent. The main theoretical point is that an organism can engage in cultural learning of this type only when it understands others as intentional agents, like the self, who have a perspective on the world that can be followed into, directed, and shared. Indeed, a strong argument can be made that children can only understand a symbolic convention in the first place if they understand their communicative partner as an intentional agent with whom one may share attention—since a linguistic symbol is nothing other than a marker for an intersubjectively shared understanding of a situation (Tomasello 1995, 1998, 1999). As a point of comparison, children with autism do not understand other persons as intentional agents, or they do so to only an imperfect degree, and so they do poorly at imitative learning of intentional actions in general (Smith & Bryson 1994): Only half of them ever learn any language at all, and those who do learn some language do poorly in word-learning situations such as those just described (Baron-Cohen et al 1997). As we see below, nonhuman primates are not very human-like in these kinds of social-cognitive and cultural learning skills either.

Linguistic Symbols and Cognitive Representation

One of the most interesting things about the process of language acquisition is that the adults from whom the child is learning went through the same process earlier in their lives, and across generations the symbolic artifacts that comprise English or Turkish, or whatever language, accumulate modifications as new linguistic forms are created by grammaticization and other processes of language change (e.g. Traugott & Heine 1991). Thus, today’s child is learning the whole historically derived conglomeration. Consequently, when the child learns the conventional use of these well-traveled symbols, what she is learning is the ways her forbears in the culture found useful for manipulating the attention of others in the past. And because the peoples of a culture, as they move through historical time, evolve many and varied purposes for manipulating the attention of one another (and because they need to do this in many different types of discourse situations), today’s child is faced with a whole panoply of different linguistic symbols and constructions that embody many different attentional construals of any given situation. As a sampling, languages embody attentional construals based on the following (for more specifics, see Langacker 1987): generality-specificity ("thing," “furniture,” “chair,” “desk chair”); perspective ("chase-flee,” “buy-sell,” “come-go,” “borrow-lend”); function ("father,” “lawyer,” “man,” “American”) ("coast,” “shore,” “beach”). There are many more specific perspectives that arise in grammatical combinations of various sorts: “She smashed the vase” versus “The vase was smashed.” It is at about 18 months of age that children first begin to predicate multiple things about objects to which they and the adult are jointly attending, for example by saying that this ball is either “wet” or “big” or “mine”—all about one and the same object (Reed 1995; Tomasello 1988, 1995).

Consequently, as a young child internalizes a linguistic symbol, as she learns the human perspective embodied in that symbol, she cognitively represents not just perceptual or motoric aspects of a situation but also one way, among other
ways of which she is also aware, that the current situation may be attentionally construed by “us,” the users of the symbol. The way that human beings use linguistic symbols thus creates a clear break with straightforward perceptual or sensory-motor cognitive representations. It is true that a prelinguistic child, or a nonhuman primate, may construe situations in more than one way: One time a conspecific is a friend and the next time an enemy; one time a tree is for climbing to avoid predators and the next time it is for making a nest in. In these different interactions with the same entity, the individual is deploying its attention differentially, depending on its goal at that moment. But shifting attention sequentially in this manner as a function of goal is not the same thing as knowing simultaneously a number of different ways in which something might be construed—in effect, imagining at the same time a number of different possible goals and their implications for attention. An individual language user looks at a tree and, before drawing the attention of her interlocutor to that tree, must decide, based on her assessment of the listener’s current knowledge and expectations, whether to say “that tree over there,” “it,” “the oak,” “that hundred-year-old oak,” “the tree,” “The bagswing tree,” “that thing in the front yard,” “the ornament,” “the embarrassment,” or any of a number of other expressions. She must decide whether the tree is in, is standing in, is growing in, was placed in, or is flourishing in the front yard. And these decisions are not made on the basis of the speaker’s direct goal with respect to the object or activity involved, but rather they are made on the basis of her goal with respect to the listener’s interest and attention to that object or activity. This means that the speaker knows that the listener shares with her these same choices for construal—again, all available simultaneously. Indeed, the fact is that while she is speaking, the speaker is constantly monitoring the listener’s attentional status (and vice versa), which means that both participants in a conversation are always aware that there are at least their two actual perspectives on a situation, as well as many more that are symbolizable in currently unused symbols and constructions.

The point is not just that linguistic symbols provide handy tags for human concepts, or even that they influence or determine the shape of those concepts, though they do both of these things. The point is that the intersubjectivity of human linguistic symbols—and their perspectival nature as one offshoot of this intersubjectivity—means that linguistic symbols do not represent the world directly, in the manner of perceptual or sensory-motor representations, but rather they are used by people to induce others to construe certain perceptual/conceptual situations—to attend to them—in one way rather than in another. This breaks symbols away from the sensory-motor world of objects in space and puts them instead in the realm of the human ability to view the world in whatever way is convenient for the communicative purpose at hand. The most important point in the current context is that as children participate in these communicative exchanges, they internalize, in something like the way Vygotsky (1978) envisioned the process, the perspectives of other persons. The internalization process is not something mystical or unanalyzable, as it is sometimes characterized, but rather it is just the normal process of imitative learning when linguistic (and perhaps other communicative)
conventions are involved. That is, in imitatively learning a linguistic symbol from an adult, the child comprehends that by using a particular symbol she intends for another to pay attention to some specific aspect of their shared experience. When the child attempts to appropriate the use of this communicative convention for her own use, she must reverse roles: If others wish her to focus on this same aspect of reality, they must use that same symbol toward her. This learning process is indeed what creates the communicative convention in the first place, in the sense that it initiates the child into the convention. Because linguistic symbols are perspectival, i.e. used to focus the attention of others on specific aspects of situations as opposed to other aspects, if the child is to use the symbol in its conventionally appropriate manner she must understand something of the adult’s perspective. It is in this sense and only in this sense that internalization involves a special form of social learning—cultural learning—in which the child internalizes the perspective of another person.

Some of the effects of operating with symbols of this type are obvious, in terms of flexibility and relative freedom from perception. But some are more far reaching and quite unexpected, in the sense that they give children truly new ways of conceptualizing things, such as treating objects as actions (“he porched the newspaper”), actions as objects (“skiing is fun”), and all kinds of metaphorical construals of things (“love is a journey”) (Lakoff & Johnson 1980, Lakoff 1987, Johnson 1987). These new ways of conceptualizing and thinking result from the accumulated effects of engaging in linguistic communication with other persons for some years during early cognitive development. More extended bouts of discourse interaction with other persons also create opportunities for explicitly exploring and comparing differing verbally expressed perspectives on situations. Perhaps of special importance are discourse interactions in which the communicative partner provides a verbally expressed perspective on the child’s previous verbally expressed perspective, since in this case the internalization of the other’s perspective helps to create children’s ability to self-regulate, self-monitor, and reflect on their own cognition (Vygotsky 1978).

**NONHUMAN PRIMATE CULTURE**

McGrew (1998) claims that nonhuman primates engage in social activities that are best characterized as cultural in that they share all the essential features of human culture. I agree with this (Tomasello, 1990, 1994, 1996a). Nevertheless, at the same time I insist that human culture has, in addition, some unique characteristics (as may the cultures of other primate species). The most important of these, at the macro-level, is the fact that many human cultural traditions and artifacts accumulate modifications over time, whereas this does not seem to be the case for nonhuman primate cultural traditions (Tomasello et al 1993a, Tomasello & Call 1997). The reason for this difference resides in the micro-level processes by which individuals learn things from and through one another, and as just elaborated, it may be the case that in their early ontogenies, human beings do this in some species-unique ways. To see whether these skills are indeed unique to
humans, we should look briefly at nonhuman primate culture, nonhuman primate social learning, and nonhuman primate social cognition.

**Japanese Macaque Potato Washing**

The most often-cited case of nonhuman primate culture is the case of Japanese macaque potato washing (Kawamura 1959, Kawai 1965). In 1953, an 18-month-old female named Imo was observed to take pieces of sweet potato, given to her and the rest of the troop by researchers, and to wash the sand off of them in some nearby water (at first a stream and then the ocean). About 3 months after she began to wash her potatoes, the practice was observed in Imo’s mother and two of her playmates (and then their mothers). During the next 2 years, seven other youngsters also began to wash potatoes, and within 3 years of Imo’s first potato washing approximately 40% of the troop was doing the same. It was thought significant that it was Imo’s close associates who learned the behavior first, and their associates directly after, in that it suggested that the means of propagation of this behavior was some form of imitation, in which one individual actually copied the behavior of another.

The interpretation of these observations in terms of culture and imitation has two main problems, however. The first is that potato washing is much less unusual a behavior for monkeys than was originally thought. Brushing sand off food turns out to be something that many monkeys do naturally, and indeed this had been observed in the Koshima monkeys prior to the emergence of washing. It is thus not surprising that potato washing was also observed in four other troops of human-provisioned Japanese macaques soon after the Koshima observations (Kawai 1965), which implies that at least four individuals learned on their own. Also, in captivity, individuals of other monkey species learn quite rapidly on their own to wash their food when provided with sandy fruits and bowls of water (Visalberghi & Fragaszy 1990). The second problem has to do with the pattern of the spread of potato washing behavior within the group. The spread of the behavior was relatively slow, with an average time of over 2 years for acquisition by all the members of the group who learned it (Galef 1992). Moreover, the rate of spread did not increase as the number of users increased. If the mechanism of transmission was imitation, an increase in the rate of propagation would be expected as more demonstrators became available for observation over time. In contrast, if processes of individual learning were at work, a slower and steadier rate of spread would be expected—which was in fact observed. The fact that Imo’s friends and relatives were first to learn the behavior may be due to the fact that friends and relative stay close to one another, and thus Imo’s friends very likely went near the water more often during feeding than other group members, increasing their chances for individual discovery.

**Chimpanzee Tool Use**

Perhaps the best examples to examine in the current context are humans’ closest primate relatives, the chimpanzees—especially with regard to tool use and ges-
tural communication, the two behaviors for which cultural transmission has been most often claimed. First, there are a number of population-specific tool use traditions that have been documented for different chimpanzee communities, for example termite-fishing, ant-fishing, ant-dipping, nut-cracking, and leaf-sponging (for a review, see McGrew 1992). Sometimes the “same” tradition even shows variability between groups. For instance, members of the Kasakela community at Gombe (as well as some other groups elsewhere) fish for termites by probing termite mounds with small, thin sticks, whereas in other parts of Africa there are chimpanzees who perforate termite mounds with large sticks and attempt to scoop up the insects by the handful. One possible explanation is that the chimpanzees in the western parts of Africa are able to destroy termite mounds with large sticks because the mounds are soft from much rain, whereas in the east they cannot use this strategy because the mounds are too hard. Each individual thus reinvents the wheel for itself, with population differences due to the different local ecologies of the different groups—so-called environmental shaping.

Although environmental shaping is likely a part of the explanations for group differences of behavior for all species, experimental studies have demonstrated that more than this is going on in chimpanzee culture (see also Boesch et al 1994). Tomasello (1996a) reviewed all the experimental evidence on chimpanzee imitative learning of tool use (a total of five studies) and concluded that chimpanzees are good at learning about the dynamic affordances of objects they see being manipulated by others, but they are not skillful at learning from others a new behavioral strategy per se. For example, if a mother rolls over a log and eats the insects underneath, her child will likely follow suit. This is simply because the child learned from the mother’s act that there are insects under the log—a fact she did not know and very likely would not have discovered on her own. But she did not learn how to roll over a log or to eat insects; these are things she already knew how to do or could learn how to do on her own. (Thus, the youngster would have learned the same thing if the wind, rather than its mother, had caused the log to roll over and expose the ants.) This is what has been called emulation learning because it is learning that focuses on the environmental events involved—the results or changes of state in the environment that the other produced—rather than on the actions that produced those results (Tomasello 1990, 1996a).

Emulation learning is a very intelligent and creative learning process that, in some circumstances, is a more adaptive strategy than imitative learning. For example, Nagell et al (1993) presented chimpanzees and 2-year-old human children with a rake-like tool and an out-of-reach object. The tool was such that it could be used in either of two ways, leading to the same end-result of obtaining the object. For each species, one group of subjects observed a demonstrator employ one method of tool use (less efficient) and another group of subjects observed another method of tool use (more efficient). The result was that whereas human children in general copied the method of the demonstrator in each of the two observation conditions (imitative learning), chimpanzees used the same method or methods to obtain the object no matter which demonstration they observed ( emulation learning). The interesting point is that many children insisted
on this reproduction of adult behavior even in the case of the less-efficient method, leading to less successful performance than the chimpanzees in this condition. Imitative learning is not a “higher” or “more intelligent” learning strategy than emulation learning; it is simply a more social strategy—which, in some circumstances and for some behaviors, has some advantages. This emulation learning explanation also applies to other studies of chimpanzee social learning of tool use activities, such as those by Whiten et al (1996) and Russon & Galdikas (1993).

Chimpanzees are intelligent and creative in using tools and understanding changes in the environment brought about by the tool use of others, but they do not seem to understand the instrumental behavior of conspecifics the same way as do humans. For humans, the goal or intention of the demonstrator is a central part of what they perceive, and indeed the goal is understood as something separate from the various behavioral means that may be used to accomplish the goal. An observer’s ability to separate goal and means serves to highlight for herself the demonstrator’s method or strategy of tool use as an independent entity—the behavior she is using in an attempt to accomplish the goal, given the possibility of other means of accomplishing it. In the absence of this ability to understand goal and behavioral means as separable in the actions of others, chimpanzee observers focus on the changes of state (including changes of spatial position) of the objects involved during the demonstration, with the motions of the demonstrator being, in effect, just other motions. The intentional states of the demonstrator, and thus her behavioral methods as distinct behavioral entities, are simply not a part of their experience.

Chimpanzee Gestural Communication

The other well-known case is the gestural communication of chimpanzees, for which there are also some population-specific behaviors (Goodall 1986, Tomasello 1990). There is one reasonably well-documented example from the wild. Nishida (1980) reported “leaf-clipping” in the Mahale K group of chimpanzees—thought to be unique to that group but later observed by Sugiyama (1981) in another group across the continent. The reporting of data for individuals in these studies showed that there were marked individual differences within the groups in how (toward what end) the signal was used, for example for sexual solicitation, aggression toward groupmates, or aggression toward humans. One hypothesis is that after one individual used leaf-clipping to make noise (the tearing of the rigid dead leaves makes a very loud noise), others learned via emulation to make the same noise (i.e. they learned the affordances of the leaf). This had different attention-getting effects on conspecifics in the different groups, however, and these were then learned as contingencies. The fact that leaf-clipping has been observed in more than one group, who have had no opportunity to observe one another, raises the possibility of some such process.

This possibility is also supported by studies with captive chimpanzees. In ongoing studies of the gestural signaling of a captive colony of chimpanzees, Tomasello and colleagues have asked whether youngsters acquire their gestural
signals by imitative learning or by a process of ontogenetic ritualization (Tomasello et al. 1985, 1989, 1994, 1997). In ontogenetic ritualization, a communicatory signal is created by two organisms shaping each other’s behavior in repeated instances of a social interaction. For example, an infant may initiate nursing by going directly for the mother’s nipple, perhaps grabbing and moving her arm in the process. In some future encounter the mother might anticipate the infant’s impending behavioral efforts at the first touch of her arm, and so become receptive at that point, leading the infant on some future occasion still to abbreviate its behavior to a touch on the arm while waiting for a response (“arm-touch” as a so-called intention movement). Note that there is no hint here that one individual is seeking to reproduce the behavior of another; there is only reciprocal social interaction over repeated encounters that results eventually in a communicative signal. This is presumably the way that most human infants learn the “arms-overhead” gesture to request that adults pick them up (Lock 1978).

All of the available evidence suggests that ontogenetic ritualization, not imitative learning, is responsible for chimpanzees’ acquisition of communicative gestures. First, there are a number of idiosyncratic signals that are used by only one individual (see also Goodall 1986). These signals could not have been learned by imitative processes and so must have been individually invented and ritualized. Second, longitudinal analyses have revealed quite clearly, by both qualitative and quantitative comparisons, that there is much individuality in the use of gestures, with much individual variability both within and across generations, which suggesting something other than imitative learning. It is also important that the gestures that are shared by many youngsters are gestures that are also used frequently by captive youngsters raised in peer groups with no opportunity to observe older conspecifics. Finally, in an experimental study, Tomasello et al. (1997) removed an individual from the group and taught her two different arbitrary signals by means of which she obtained desired food from a human. When she was then returned to the group and used these same gestures to obtain food from a human in full view of other group members, there was not one instance of another individual reproducing either of the new gestures.

The clear conclusion is that chimpanzee youngsters acquire the majority, if not the totality, of their gestures by individually ritualizing them with one another. The explanation for this learning process is analogous to the explanation for emulation learning in the case of tool use. Like emulation learning, ontogenetic ritualization does not require individuals to understand the behavior of others as separable into means and goals in the same way as does imitative learning. Imitatively learning an arm-touch as a solicitation for nursing would require that an infant observe another infant using an arm-touch and know what goal it was pursuing (viz. nursing), so that when it had the same goal it could use the same behavioral means (viz. arm-touch). Ritualizing an arm-touch, on the other hand, only requires the infant to anticipate the future behavior of a conspecific in a context in which it (the infant) already has the goal of nursing. Ontogenetic ritualization is thus, like emulation learning, an intelligent and creative social learning process that is important in all social species, including humans. But it is not a learning
process by means of which individuals attempt to reproduce the behavioral strategies of others.

Nonhuman Primate Social Learning and Social Cognition

Chimpanzee tool use and gestural communication thus provide us with two very different sources of evidence about nonhuman primate social learning. In the case of tool use, it is likely that chimpanzees acquire the tool use skills they are exposed to by a process of emulation learning. In the case of gestural signals, it is likely that they acquire their communicative gestures through a process of ontogenetic ritualization. Both emulation learning and ontogenetic ritualization require skills of cognition and social learning, each in its own way, but neither requires skills of imitative learning, in which the learner comprehends both the demonstrator’s goal and the strategy she is using to pursue that goal—and then in some way aligns this goal and strategy with her own. Indeed, emulation learning and ontogenetic ritualization are precisely the kinds of social learning one would expect of organisms that are intelligent and quick to learn, but that do not understand others as intentional agents with whom they can align themselves.

The other main process involved in cultural transmission as traditionally defined is teaching. Whereas social learning comes from the “bottom up,” as ignorant or unskilled individuals seek to become more knowledgeable or skilled, teaching comes from the “top down,” as knowledgeable or skilled individuals seek to impart knowledge or skills to others. The problem in this case is that there are few systematic studies of teaching in nonhuman primates. The most thorough study is that of Boesch (1991), in which chimpanzee mothers and infants were observed in the context of tool use (nut cracking). Boesch discovered that a mother does a number of things that serve to facilitate the infant’s activities with the tool and nuts, such as leaving the tools idle while she goes to gather more nuts (which she would not do if another adult were present). But the interpretation of the mother’s intention in such cases is far from straightforward. Moreover, in the category of “active instruction,” in which the mother appears to be actively attempting to instruct her child, Boesch observed only two possible instances (over many years of observation). These two instances are also difficult to interpret, in the sense that the mother may or may not have had the goal of helping the youngster learn to use the tool. On the other hand, although there is much variability across different human societies, adult humans in all cultures actively instruct their young on a regular basis in one way or another (Kruger & Tomasello 1996). Along with imitative learning, the process of active instruction is likely crucial to the uniquely human pattern of cumulative cultural evolution as well.

It should be acknowledged, of course, that this way of viewing nonhuman primate culture and social learning is not the only way of viewing them, and indeed researchers such as McGrew (1998) and Boesch (1999) would likely disagree with many of the current conclusions. Indeed, there are also a number of widely publicized studies purportedly demonstrating that nonhuman primates have theories of mind, deceive their conspecifics, and engage in many other kinds of “mind-reading” activities (e.g. Byrne 1995). If these studies represented the true picture
of nonhuman primate social cognition, it would be a complete mystery why they did not engage in more powerful forms of cultural learning. However, there are in actuality surprisingly few of these studies, and as in all scientific paradigms, a few studies is not enough to settle opinion. My own interpretation of these studies is that they represent cases in which nonhuman primates have acquired clever strategies to manipulate the behavior, not the mental states, of others (for detailed analyses of all the relevant studies, see Tomasello 1996b; Tomasello & Call 1994, 1997). If they are not dealing with mental or intentional states, but only behavior, nonhuman primates will not engage in the kinds of cultural learning that lead to the ratchet effect—and that thus create human-like cultural niches full of symbolic artifacts and instructional formats for the cognitive ontogeny of their offspring.

With regard to nonhuman primate culture in particular, Boyd & Richerson (1996), in a paper entitled “Why culture is common, but cultural evolution is rare,” hypothesize that humans and other animals both engage in the same kinds of social and imitative learning but with a quantitative difference. Thus, chimpanzees may have some imitative learning abilities, but they may display them less consistently or in a narrower range of contexts than do humans—or it may even be that only some individuals have these skills. These authors go on to make a quantitatively based evolutionary argument that this rarity itself can make cultural evolution of the cumulative type impossible. So it might be that a quantitative difference in social learning leads to a qualitative difference in the nature of cultural traditions, and in particular how they change and evolve over time. Although there are currently no easy ways to quantify the frequency of imitative learning in the societies of different species, this hypothesis is intriguing because it posits the existence of the kind of variation that might have characterized the human-chimpanzee common ancestor and with which subsequent evolutionary processes might have worked.

**Enculturated Apes**

It may be objected that there are a number of convincing observations of chimpanzee imitation in the literature, and indeed there are a few. It is interesting, however, that basically all the clear cases in the exhaustive review of Whiten & Ham (1992) concern chimpanzees that have had extensive amounts of human contact. In many cases, this has taken the form of intentional instruction involving human encouragement of behavior and attention, and even direct reinforcement for imitation for many months [e.g. 7 months of training in the case of Hayes & Hayes (1952) and 4 months of training in the case of Whiten & Custance (1996)]. This raises the possibility that imitative learning skills may be influenced, or even enabled, by certain kinds of social interaction during early ontogeny.

Confirmation for this point of view is provided in a study by Tomasello et al (1993b). It compared the imitative learning abilities of mother-reared captive chimpanzees, enculturated chimpanzees (raised like human children and exposed to a language-like system of communication), and 2-year-old human children. Each subject was shown 24 different and novel actions on objects, and each sub-
ject’s behavior on each trial was scored as to whether it successfully reproduced (a) the end result of the demonstrated action and/or (b) the behavioral means used by the demonstrator. The major result was that the mother-reared chimpanzees reproduced both the end and the means of the novel actions (i.e. imitatively learned them) hardly at all. In contrast, the enculturated chimpanzees and the human children imitatively learned the novel actions much more frequently, and they did not differ from one another in this learning. Interesting corroboration for this latter finding is the fact that earlier in their ontogeny, these same enculturated chimpanzees seemed to learn many of their human-like symbols by means of imitative learning (Savage-Rumbaugh 1990).

For the issue of chimpanzee culture in the wild, these results raise an important question. Which group of captive chimpanzees is more representative of chimpanzees in their natural habitats: mother-reared or enculturated? Are enculturated chimpanzees simply displaying more species-typical imitative learning skills because their more enriched rearing conditions more closely resemble those of wild chimpanzees than do the impoverished rearing conditions of other captive chimpanzees? Or might it be the case that the human-like socialization process experienced by enculturated chimpanzees differs significantly from the natural state and, in effect, helps to create a set of species-atypical abilities more similar to those of humans? There can be no definitive answer to these questions at this time, but one possibility is that a human-like sociocultural environment is an essential component in the development of human-like social-cognitive and imitative learning skills, no matter the species. That is, this is true not only for chimpanzees but also for human beings—a human child raised in an environment lacking intentional interactions and other cultural characteristics in all likelihood would also not develop human-like skills of imitative learning.

The hypothesis is thus that the understanding of the intentions of others, necessary for reproducing another’s behavioral strategies, develops in, and only in, the context of certain kinds of intentional interactions with others (Tomasello et al. 1993a). More specifically, to come to understand others in terms of their intentions requires that the learner him- or herself be treated as an intentional agent in which another organism encourages attention to and specific behaviors toward some object of mutual interest—often reinforcing in some manner the learner’s successful attempts in this direction (Call & Tomasello 1994, 1996). Such interactions are not sufficient, of course, as many animals are subjected to all kinds of human interaction, and even direct instruction, without developing human-like skills of imitative learning (and the same is true of human children with autism). The important point for current purposes is that in terms of these dimensions of social interaction, captive chimpanzees raised by conspecifics are a better model for wild chimpanzees than are chimpanzees raised in human-like cultural environments—since wild chimpanzees receive little in the way of direct instruction from conspecifics.

A corollary hypothesis is thus that the learning skills that chimpanzees develop in the wild in the absence of human interaction (i.e. skills involving individual learning supplemented by emulation learning and ritualization) are sufficient to
create and maintain their species-typical cultural activities, but they are not sufficient to create and maintain human-like cultural activities displaying the ratchet effect and cumulative cultural evolution. The fact that chimpanzees and other great apes raised from an early age and for many years in human-like cultural environments may develop some aspects of human social cognition and cultural learning demonstrates the power of cultural processes in ontogeny in a particularly dramatic way. The effect of cultural environments on nonhuman primate cognitive development is thus a question that deserves more empirical investigation.

CONCLUSIONS

It is easy to observe a human behavior and posit a specific gene for that behavior, with no research into the genetics of the situation. Many scholars and popularists of biological and social sciences make their living doing just that. But when we have behaviors that are unique to a species, we have serious time constraints on such hypotheses, and so the positing of a large number of significant genetic events becomes highly implausible. In the case of humans, the time frame for the emergence of their unique cognitive skills is almost certainly in the range of 2 million to 0.3 million years ago—with my own theoretical bias being toward the smaller of those figures. But the genetic event that happened at that time was not an everyday genetic event. It was not an everyday genetic event because it did not just change one relatively isolated characteristic, it changed the nature of primate social cognition, which changed the social-cultural transmission process characteristic of primates, which led to a series of cascading sociological and psychological events in historical time. The new form of social cognition that started the entire process involved understanding other persons as intentional agents like the self, and the new process of social-cultural transmission involved several forms of cultural learning, the first and most important being imitative learning. These new forms of cultural learning created the possibility of a kind of ratchet effect in which human beings not only pooled their cognitive resources contemporaneously, they also built on one another’s cognitive inventions over time. This new form of cultural evolution thus created artifacts and social practices with a history, so that each new generation of children grew up in something like the accumulated wisdom of their entire social group, past and present.

And so, while not denying that a significant genetic event happened in human cognitive evolution, probably fairly recently, I would deny that this event specified the detailed outcomes we see in adult humans today. In my view, that genetic event merely opened the way for some new social and cultural processes that then, with no further genetic events, created many, if not all, of the most interesting and distinctive characteristics of human cognition. Perhaps of special importance was a new form of perspectively based cognitive representation that emerged when children began to learn and use linguistic symbols, evolved over historical time for inducing others to construe certain situations in certain ways, which necessitated an internalization and representation of the different perspec-
tives of other persons. This story is not as simple as the “genes create all novelties” story, but it accords better with data from both anthropological and developmental psychological investigations. Modern human adult cognition is a result not just of processes of biological evolution, but also of other processes that human biological evolution made possible in both cultural historical time and individual ontogenetic time.


LITERATURE CITED


Gibson E, Rader N. 1979. Attention: the perceiver as performer. In Attention and Cog-
Savage-Rumbaugh S. 1990. Language as a cause-effect communication system. Philos. Psychol. 3:55–76
Tomasello M. 1988. The role of joint atten—


