The Archaeology of Perception

Traces of Depiction and Language

by Iain Davidson and William Noble

Depiction, particularly the making of images to resemble things, can only have emerged prehistorically in communities with shared systems of meanings. We argue, on the basis of an articulation of Gibson's ecological theory of perception, Mead's distinction between communication and language, and a portmanteau theory of language and mind relying on the insights of, among others, Ryle, Vygotsky, and Olson, that depiction transforms communication into language. The rapid change in numerous practices observable at the end of the Upper Pleistocene becomes understandable when communication is seen to be turning into language as here defined. It is for this reason that the period in question represents the point of evolution of modern human beings.


The present paper was submitted in final form 23 VIII 88.

Modern human beings are uniquely characterized by reflective language—the business of accounting for, describing, avowing, gossiping, and complaining about what they see, hear, taste, touch, and smell. How did this capacity—the power to talk about what is perceived—evolve? The answer lies in the equally unique human practice of depiction—the making of images by various means in various media. Recent papers have addressed the problem of the origins of image making (Davis 1986a, Halverson 1987, Chase and Dibble 1987). We argue that communication of some sort is necessary for depiction and, further, that depiction transforms communication into language. A picture of a bison is not a bison, and the word "bison" is neither a picture of a bison nor a bison. Both the picture and the word represent reality. Language cannot resemble what it represents except through the occasional onomatopoeic word; depiction can, and thus stands halfway between reality and language. The argument we are presenting is about how hominids came to use things to stand for other things: how hominids became human.

The concepts critical to our argument are mimicry, depiction, and reflective language. By "mimicry" we mean the deliberate copying, by means of bodily posture or gesture, of features of the environment. We will draw particular attention to the mimicking of features of an animal's form or gait. By "depiction" we mean the deliberate making of marks on surfaces or other modification of objects such that an image or pattern results which is recognizable as an image of something (or is recognizable as a pattern). Depiction is thus distinct not only from the traces left by a passing animal (footprints, etc.) but, indeed, from the earliest hominid markings, such as those discussed by Marshack (1976), which have no discernible form or pattern.

We have used the term "deliberate" in defining both these concepts. In contemporary human conduct, everyone knows what it means to do something deliberately—for example, with intention, or with an aim in mind. There is less certainty about how to apply the notion of "deliberateness" to the conduct of hominids in the course of their evolution into human beings. This is because "deliberateness" in contemporary humans is very much tied in with their capacity to represent to themselves, by means of language, their aims and intentions. As we are not at all sure that hominids had the equivalent of modern human language, we must be more circumspect about their self-representational ability. "Mimicry" and "depiction," even in their ancient practice, are deliberate in that they are not to be understood as the results of accident or incident. Something imitative might initially have occurred accidentally or incidentally, thereafter developing into the deliberate practice we now call mimicry and depiction.

Defining "reflective language" introduces the substance of our argument. The word "reflective" is used for emphasis; language is by definition reflective in that it is bound in with the capacity to adopt an attitude of reference to the perceived world (Berger and Luckmann 1967). Such an attitude can be more clearly appreciated.
when considered in contrast to an attitude of monitoring. For a creature to monitor its environment entails its being alert to changes that might or do affect it. All animate life monitors its environment in one or more ways, with varying degrees of activeness (mobility) and to varying extents of feature differentiation. In contrast to monitoring, reference to what is perceived entails being able to take notice of one’s monitoring. It entails reflection upon what is perceived, not simply the registering of environmental events (Harré and Secord 1972). Such reflection is rendered by means of language. It is through language that we “realize” what we perceive. Confusion has resulted from the assumption that language merely sits on top of some more generalized capacity for consciousness about “what’s going on”—a sort of wordless awareness that is uniform among higher mammals and simply happens to be shattered about by the human animal. It is, however, incoherent to assume such awareness-in-common in creatures about which nothing can be understood with respect to their “phenomenal” universe because nothing can be exchanged with them via a system of shared meanings (Descartes 1985 [1637], Coulter 1983).

Looking for the Origins of Language

There have been various approaches to the origins of language, a topic whose history is ancient (see Hewes 1976, Aarsleff 1976). The present-day invisibility of early language means that conventional archaeological methods of investigation are not sufficient. As a result there have been a number of arguments that we consider fallacious.

The commonest fallacy is to assume that the development of children is a literal recapitulation of the evolutionary stages of language (see Parker 1985, Rumsey 1985). Such approaches may marry a speculative sequence of stages in the evolution of language with a speculative interpretation of the abilities of fossil hominids, gaining no strength in fact from the combination.

A more empirical but equally fallacious approach is to identify the existence of structures in the fossils seemingly similar to those related to speech in modern humans and to suggest that this entails some language abilities (e.g., Tobias 1981, Falk 1983). Our argument will be that such structures and capacities emerged before functions (see Falk 1985) such as spoken language. Vocalizing and the muscular control necessary for it may have evolved closely with the neural structures which are still associated with them. But vocalizing is not equivalent to language. Falk was close to the point in saying, “we should not ask whether australopithecines could sing, or sign, instead we should ask what it was they had to talk about” (Falk 1980a). Talking about things entails reflection, and that cannot be found in the structures of the brain.

Then there are the clever-chimpanzee fallacies, which point to the similarities in genetics between chimpanzees and modern people (see Jorde 1985), to similarities between the abilities of chimpanzees in the wild and the evidence from the Lower Pleistocene (see table 1) archaeological record (see, e.g., McGrew 1986), and to the most generous interpretations of the results of attempts to teach either American Sign Language or some concocted pattern of symbolic communication to chimpanzees or gorillas (see, e.g., summary in Goodall 1986:15–42). Such studies allow the belief that these characters indicate a capacity for language in chimpanzees, which thus must have been shared with our common ancestor. In this interpretation australopithecines would have derived new behavioral characters in the direction of becoming human and thus have been even more capable of language. The fallacy arises because some of the observed behaviors in the wild (e.g., Kortlandt 1986) and, of course, outcomes from all of the chimpanzee language experiments are plausibly due to imitation of modern humans (Seidenberg and Petitto 1979) and therefore not directly applicable to discussion of the abilities of hominids.

None of these fallacies really addresses the evidence of the archaeological record of hominid behavior.

The approach to the archaeological record is best expressed by Isaac (1976). He suggested that there are two
TABLE 2
Documented Early Objects with Deliberate Marks

<table>
<thead>
<tr>
<th>Site</th>
<th>Object</th>
<th>Industry</th>
<th>Date</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>La Quina</td>
<td>Drilled fox tooth</td>
<td>Mousterian</td>
<td></td>
<td>Marshack [1976], Martin [1909]</td>
</tr>
<tr>
<td>La Quina</td>
<td>Pierced reindeer phalanx</td>
<td>Mousterian</td>
<td></td>
<td>Marshack [1976], Martin [1909]</td>
</tr>
<tr>
<td>La Ferrassie</td>
<td>Bone with fine incised parallel lines</td>
<td>Mousterian</td>
<td></td>
<td>Marshack [1976], Capitan and Peyrony [1921]</td>
</tr>
<tr>
<td>Bacho Kiro</td>
<td>Bone with engraved zigzags</td>
<td>Mousterian</td>
<td>&gt;43,000</td>
<td>Marshack [1976], Kozłowski [1982]</td>
</tr>
<tr>
<td>Tata</td>
<td>Mammoth tooth section, modified</td>
<td>Mousterian</td>
<td>116–78,000</td>
<td>Schwarz and Skolek [1982]</td>
</tr>
<tr>
<td>Pech de l’Azé</td>
<td>Rib fragment with incised lines</td>
<td>Acheulean</td>
<td>123–103,000</td>
<td>Schwarz and Blackwell [1983]</td>
</tr>
<tr>
<td>Pech de l’Azé</td>
<td>Pierced bone fragment</td>
<td>Mousterian</td>
<td></td>
<td>Bordes [1969]</td>
</tr>
<tr>
<td>Morin</td>
<td>Bone fragment with incised lines</td>
<td>Mousterian</td>
<td></td>
<td>Freeman [1983]</td>
</tr>
<tr>
<td>Sclayn</td>
<td>Bear tooth with grooves</td>
<td>Middle Paleolithic</td>
<td></td>
<td>Chase and Dibble [1987]</td>
</tr>
<tr>
<td>Berekhat Ram</td>
<td>?Figurine</td>
<td>Acheulean</td>
<td>&gt;230,000</td>
<td>Goren-Inbar [1986]</td>
</tr>
<tr>
<td>Kebara</td>
<td>Incised bones</td>
<td>Mousterian</td>
<td></td>
<td>Davis [1974]</td>
</tr>
<tr>
<td>Klasies River Mouth</td>
<td>Bone with parallel grooves</td>
<td>Middle Stone Age</td>
<td></td>
<td>Singer and Wymer [1982]</td>
</tr>
<tr>
<td>Klasies River Mouth</td>
<td>Ribs with serrated edges</td>
<td>Middle Stone Age</td>
<td></td>
<td>Singer and Wymer [1982]</td>
</tr>
</tbody>
</table>

attitudes to the evidence for the rate of change of material culture. In both, the record of stone tools is seen to have a long period [from about 1.6 to 0.2 million years] of relative stagnation [see Isaac 1972], when the stone industries throughout the world [Huang 1987] were dominated by the production of flakes and handaxes, followed by a period of increasingly rapid change. Isaac distinguished between those who see this history as an exponential curve indicating continuous change and those who suggest that there is a break in the curve which represents a punctuation in the abilities of hominids. Gовlett’s [1984a, b] gradualistic approach is in the former camp; we suggest the latter.

Gовlett [1984a, b; see also Oakley 1954] proposed that the evidence of stone tool making suggests such complex behavior on the part of early hominids that we must suppose that the skills were transmitted by language [Gовlett 1984a:55]. Against this it can be said that the repetitive patterning which we see in handaxes would have allowed learning of the skills of stone tool making by imitation. This would not seem to have required language.

Although Marshack [e.g., 1976] has been one of the pioneers in recognizing the link between the archaeological evidence for the origins of depiction and the origins of language—the attitude which emphasizes the break in the curve of innovation—his work also exemplifies the second archaeologists’ fallacy, assuming the abilities that one is trying to discover. Thus in discussing the earliest then-known object that was thought to be evidence of symbolic expression by hominids, the engraved bone from Pech de l’Azé [Bordes 1969], Marshack [1976:303] described the marks as in “precisely” the same “tradition” as images produced in the Upper Paleolithic. We argue that it could not have been precisely the same tradition, since tradition requires shared meaning, and meaning in the earliest marks cannot be attributed. While we are wary of his terminology, we agree with Davis [1986b: 516] that “the concept did not precede ‘the image’... for the mind [hand, eye, and brain] does not have a ‘concept’ of that which it represents until it has actually succeeded in obtaining a Representation.”

Some of these problems have arisen because archaeologists or paleoanthropologists are not familiar with the relevant aspects of social science or because other social scientists are not familiar with the evidence of archaeology and paleoanthropology. In an endeavor that, by its nature, is speculative, it may be advantageous that the two of us represent these different disciplines.

The earliest image making seems to lie around 32,000 years ago in two regions of Europe. In southwestern France there are the stone blocks from seven sites [Del Luc and Delluc 1978, 1981] with engraved signs sometimes interpreted as representing the human vulva [see Bahn 1986], and in southwestern Germany there are the figurines from Vogelherd and scratched bones from at least 25 other sites [Hahn 1972, 1983]. In addition to these clusters of repeated signs, there are scattered cases of marks on bones and “decorated” bones which are earlier in date (table 2). Marshack [1976] has argued that these can be seen as part of the origin of—that is, prior to and leading to—the Upper Paleolithic depictive tradition.

The early evidence for modified bones indicates a capacity to make things which do not have intrinsic meaning or obvious function. The Tata tooth could have been part of a system of symbolic communication, but its very uniqueness makes it unlikely that it was. Similarly, neither the marked bones from La Ferrassie and Morin nor the pierced bones from La Quina and Pech de l’Azé constitute the basis for an important use of symbols because of the lack of evidence that there was any shared meaning. On the other hand, both the so-called vulvas from France and the scratched bones which accompany the Vogelherd figurines, separately, constitute repeated
themes and ones which were restricted in period and place, so that, plausibly, meaning, whatever it was, could have been shared among observers of the objects.

Almost all of the evidence we are discussing comes from Europe and the western part of Asia, because that is the region with the largest number of known cases. It is not clear that this is necessarily because there was a single origin of such behavior in that region. An example of the alternative possibility is the evidence from Klazies River Mouth in South Africa [Singer and Wymer 1982]. The site has uncertain chronology [Binford 1984] but seems to have ochre back into the beginning of the Upper Pleistocene and three marked bones earlier than 40,000 years [Singer and Wymer 1982:116–17]. Coloured images have been found on two stone plaquettes dated to 26,000 years at Apollo 11 Cave, Namibia [Wendt 1976].

The origins of image making must be earlier than the early images such as those from Vogelherd or Apollo 11. The search for them takes us back into the early Upper Pleistocene and the Middle Paleolithic or Middle Stone Age.

The argument about the Middle Paleolithic evidence has recently been summarized by Chase and Dibble (1987) as an argument about the use of symbols. The case rests on evidence from stone artefacts, burials, other claims for ritual, including the so-called bear cult and claims for cannibalism, the presence of ochre, and other claims for “art.” None of the evidence which is said to indicate early symbolic behavior is convincing [Davidson 1987], and, except for stone tool making, we will not discuss it further. All lines of evidence point to the fundamental change in the Upper Paleolithic. In that period there is better evidence for pattern and design in stone tools, and there are more regional variations in their pattern and design. There is unequivocal burial, including multiple burial and burial with unambiguous grave goods; there is evidence for practices which can only be understood as involving ritual, and ochre was clearly used for image making. In addition, there were major alterations to behavior across the transition [Renfrew 1987:683–94].

Data which seem to relate to the origins of depiction are extremely scarce in the archaeological record, to the extent that it is unlikely that new discoveries, such as the curious pebble from Berekhat Ram [Goren-Inbar 1986], will radically alter the situation. The only way to resolve the problems they present is to develop theory which can accommodate their form and nature.

Marshack largely assumed that the early deliberately marked objects were symbolic and hence that language was already being used “to maintain and explain the tradition” [1976:307]. It is our argument here that such objects were not yet symbolic. For meaning in symbols to be shared they need to be repeatedly and widely used. Scarcity and uniqueness of form cannot be argued to indicate shared meaning. For all of these objects it would seem that there was initially no way to ensure shared meaning. They all had the capacity to be used as symbols, but we have no evidence that they were so used. We suggest that the capacity to make them symbolic emerged later than the capacity to make them.

It is not until we, in the present, are confronted by a trace whose referent is recognizable [transparent] to us (“this is a ‘bison’”), even if its significance to its maker is unknown to us [why draw a “bison”?], that we are clearly engaged with the product of a community that had an elaborated and syntactically controlled communication system. In addition, a pattern recognizable as a pattern [though referentially opaque] is in the same category as one made to be a figure and thus has the same semantic entailment. Repetition is a vital criterion for recognizing “deliberateness” of production in either case.

We make these claims on the grounds of the skills needed to execute sequences of connected acts. Hewes (1973), Bradshaw and Nettleton (1982), and Lieberman (1984) have all concluded, from reviewing the available anatomical and functional evidence, that the bodily systems enabling control of the articulators of vocal utterance are continuous with and probably successor to the systems enabling control of any fine motor sequence, such as aimed throwing [Calvin 1982], manipulation [Marzke and Shakley 1987], and manufacture. MacNeillage, Studdert-Kennedy, and Lindblom (1984) offer an account of the evolution of language and its lateralization in the brain as consequent upon the bimanual coordination required for object manipulation. And Lieberman (1984) posits “rules” allowing sequential control of complex acts such as reaching to retrieve an object for use in fulfilling an overall purpose (say, drinking from a cup). Such rules can generalize to allow control of the sequencing of utterance—its syntax.

Our claim, then, is that depiction requires sequential control of motor action [such as hominids had been incorporating through making stone tools] and must be embedded in a system of shared meanings, itself expressible through syntactically controlled utterance. The question we pose is whether, prior to the first depictions, any prehistoric utterance would have been “language,” as understood in modern terms. Whatever is recognizable as “language” can only be something akin to what modern humans do; hence we must restrict use of the term to what modern humans in fact do.

The Origins of Depiction

The resurgence of debate on the function or purpose of the earliest depiction [e.g., Stoliar 1977, 1978; Conkey 1980; Davis 1986a; Halverson 1987; Ucko 1988] is evidence of a concern of archaeologists to seek the mean-

---

2. We have avoided the term “art” at this point because it assumes a present-day perception of what was done that has no warrant.
ing(s) for the image makers of their images or their making of them. It is not our wish to join that debate. We can see no evidence that supports one or another of various plausible speculations, nor do we see how any evidence for such culture-specific meaning(s) could be obtained. Our interest, rather, is in pointing to consequences we can imagine for image makers of their engaging in this activity. In dealing with the literature in this debate, we will, thus, ignore the search for meaning(s).

Ucko [1988] persuasively dismisses most of the claims about images for which the only case for antiquity rests on assertions about stylistic development. The essence of his argument, however, is that the earliest surviving evidence may not be the first “experiments in artistic expression” (Ucko and Rosenfeld 1968:76), which may have been in perishable materials such as skin, bark, or wood. However, the existence of the small number of surviving materials from the beginning of the Upper Pleistocene, whether or not they are or were meaningful, suggests that the capacity to produce images in a non-perishable medium existed earlier than the Upper Paleolithic. Let us suppose that early Upper Pleistocene hominids created a large body of symbolic material in perishable media. We know that they were able to make marks and modify objects. It therefore would seem unlikely that, in these circumstances, such expression should never have emerged in a non-perishable medium.

It might be argued that depiction evolved as it develops during the training of a modern pictorialist. Modern humans—living in a language-rich environment—may discover this skill [within the conventions of their society [Hagen 1986]] by first making marks which bear little or no resemblance to the thing intended. Through practice, these marks come to resemble something which can be accepted as referring to an object. This cannot, however, be the path along which the ability evolved. The present-day image maker in this account has the intention to depict in a context in which it is already understood what resemblance is. Before evolving humans had the knowledge of the possibility of depiction there can have been no possibility of an intention to depict. The indecipherable precursor marks of ancient people are logically impossible as real precursors to “the recognizable.” Real indecipherable [things made purposely to be indecipherable] can only follow from the decipherable (codes transmitted en clair).

Our position is close to that of Davis [1986a:194] that “image making originated in the discovery of the representational capacity of lines, marks, or blots of color which need not and often do not have a representational status.” But he goes on to say that “by the Aurignacian period hominids had been making marks for hundreds of millennia. . . . semantic or nonsemantic marks must have been recognized as identical or similar within a recognized tolerance of variation” [p. 200]. Whereas the history may have been a long one (though probably not hundreds of millennia), there is little evidence from before the Upper Paleolithic which can be taken as representational. The marks produced could have been given meaning, by creatures like ourselves; the question is whether these were creatures like ourselves, with that capacity for recognizing the marks as marks and for giving them meaning.

Davis [1986a: 199–200] posits that, in the origin of image making, identity could be seen between marks, leading to the perception of similarity between them. Perception of identity between things allowed perception of similarity between things. This led to the possibility of perceiving similarity [although Davis posits identity] between marks and things. However, we think that a prior question needs to be addressed: how, in a world without language or at least some system of simulation, could similarities between marks (or traces) and things be perceived? There seems to be no sign, for instance, that chimpanzees can see the resemblances between pictures of things and things themselves (Winner and Ettlinger 1979). Davis has argued that there is a direct response to an image which is like “a bison” or part of a bison, once that has been brought off. The argument is that the outline of a bison does not need to be deciphered, because anyone seeing the trace and knowing bison would be able to see the resemblance between the two directly. We question the logic of this argument because it depends on the existence of modern human capacities.

The conceptual problem lies in the nature of perception and the knowledge needed for the perception of resemblance. The background from which we address this problem is a theoretical debate centered on the work of Gibson [1966, 1979] concerning whether perception is construable as direct or mediated (see Michaels and Carello 1981). Direct perception [Gibson’s theory] can be described as the sensitivity of organisms to pre-existing structures in, for example, electromagnetic or acoustic energy fluxes in the environment. Mediated-perception theories invoke cognitive abilities in the organism which construct a “working picture” of the environment, drawing on pre-established schemata, memories, expectations, and other personally possessed “data bases.” It is our view that reflective language holds the key to any resolution of this debate. Rather than rehearse this point [see Noble 1987], we lay out, briefly, the form of argument that stems from it.

Only humans possess reflective language. They are able to make self-conscious accounts of what they perceive. All living creatures perceive, in the sense that they engage in activities in their environments and with sensitivity to environmental constraints. A theory such as Gibson’s is entirely adequate to the task of explaining how all organisms (including humans for much of their everyday lives) achieve this unself-conscious coherence with respect to their environments.

Organisms that are prey for others may be discrimi-
inated in terms of the kind of movement they make or their shape [McFarland 1985:232–33]. In such predator-prey systems (toads and worms, for instance) the specification of the prey (worms) is due to its fulfillment of visual criteria, for example, wriggling, that produce predation in the other animal (the toad). A similar logic applies to avoidance responses. No mediation of perception is required to explain these events. In a real sense, the prey picks itself out of the environment. Electromagnetic energy is reflected from the worm’s body and projected to the surrounding environment, including the point of observation that happens to be occupied by the hungry toad. Human experimenters can simulate the optical effect of a wriggling worm and so reveal the feature (wriggling) to which the toad is most sensitive. It is unnecessary to invoke processes other than evolutionary to account for such events. In particular, and here we approach the problem posed by Davis’s (1986a) argument, there is no basis for reasoning that predators, in an experimental context, “see a resemblance” between simulated and actual prey. Rather, they simply respond to simulation as if confronted by the real thing.5

The human capacity to “see a resemblance,” that is to say, to take note of the fact that an appearance shares features with something else, already depends on a capacity to reflect upon what is perceived. And such reflection can only be delivered through language, or at least some system of simulation. To see a resemblance requires a capacity to represent to oneself or to others what the character of one’s perceptual experience is, or is akin to. The apparently simple achievement that Davis argues for (seeing that x resembles y) is simple only once the capacity for reflective language has emerged.

In our view, incidentally, it is the apparatus of concepts made apparent through language (“schemata,” “memories,” “expectations”) that leads theorists of mediated perception to suppose that all this must be tacitly inherent in any brain, including a toad’s.

To sum up: It is the making of things to be similar to others which provokes, or at least facilitates, the seeing of similarity between those things, or naturally occurring environmental features, and the things they resemble. Davis, it seems to us, may be using the “cognitive” result of this achievement to explain its origin.

We need to find an activity which produces simulation, such that the similarity of visible features between object and depiction can in fact begin to be seen. We believe that mimicry holds the mediating key, as against the argument from direct perception of resemblance between an object and a depiction of it (Davis 1986a). Mimicry and other forms of gestural skill can feasibly arise from tool making and use. Hewes (1973) has argued that the bodily systems engaged in the making of implements are exactly the ones employable as articulators of visible signs. More directly, the actions involved in the making and use of different tools are perceptible, hence iconic gestures for these different objects and uses can arise immediately from their forms and functions [Noble 1988]. If we postulate that gesture and posture could indicate a bison by mimicry, then a trace, as “frozen” gesture, would be directly perceptible as looking like a bison.

A hominid confronting a bison while in the presence of another hominid might have had gestural communication, developed through prior selection for increasing tool making and use, by which to indicate to the other the presence of the bison—by gesturing in the air its distinctive humped outline. Repetition of this act in a situation in which the gesture was fixed by making a trace of the outline in the mud on a cave floor would indicate “a bison” because of the mimicry that formed the gesture. This “frozen” gesture would remain to be seen in the absence of the bison. It would thus afford the materials for communication about a “bison” in the absence of a bison. Such precursory gestural communication by itself need not be considered language. By the same token, neither need vocal communication. Vocal utterance and bodily gesture are to communication what speech and sign are to language.

From Communication to Language

The act of freezing a gesture fixes the image in a place and makes it persist in time [Gibson 1966]. It thus creates memory [Vygotsky 1978 [1930]]. It allows repeatable and perceptible reference to the object by the maker and other observers. It establishes a perceptible relationship between the maker, the depiction, and the thing itself. It allows reference to two or more terms of this three-way relationship by the maker and other observers. By persisting as a sign of the gesture, it allows copying, transporting, and recopying. This is the kind of complex that starts to look more like language. It is our hypothesis that this is the way in which communication was transformed into reflective language.

Mead [1934] offers one clear analysis of the nature of language and how it contrasts with “communication” in the sense of the transmission and receipt of signals. Many animals, humans included, communicate with each other by means of vocal and other signals. The effectiveness of such signalling—which can be quite elaborately tuned to environmental circumstance (see, e.g., Seyfarth, Cheney, and Marler 1980)—is in the calling forth of an appropriate behavioral response by those who detect it (flight in the case of a warning signal, approach in the case of a courtship signal). But these signals do not constitute language, because there is no awareness that “communication” is going on. There is no reference to what is being uttered, only a response to the utterance, which is itself a response to an environmental event. Communication here is, as Mead puts it (p. 43), a “conversation of gestures.” In a dogfight, for example, the threatening attitude of one animal calls out a responding

5. Peter Freebody has pointed out that if the toad “saw the optical simulation” as merely “resembling the real thing” it would surely “not bother” to make a predatory response.
postural adjustment in the other, which in turn acts as a stimulus for the first animal, and so the chain unfolds.

Mead points to the distinction between being able to see the threatening posture of the dog and being able to see what lies behind a gesture “standing in” for such a posture [p. 45]:

In one case the observer sees that the attitude of the dog means attack, but he does not say that it means a conscious determination to attack on the part of the dog. However, if somebody shakes his fist in your face you assume that he has not only a hostile attitude but that he has some idea behind it. You assume that it means not only a possible attack, but that the individual has an idea in his experience.

He goes on to draw out the distinction implicated in these two types of events [pp. 45–46]:

When . . . that gesture [shaking of fist] means this idea behind it and it arouses that idea in the other individual, then we have a significant symbol. In the case of the dog-fight we have a gesture which calls out appropriate response; in the present case we have a symbol which answers to a meaning in the experience of the first individual and which also calls out that meaning in the second individual. Where the gesture reaches that situation it has become what we call “language.” It is now a significant symbol and it signifies a certain meaning.

Language, then, is a system of recognizable meanings arising out of shared, and thus repeated, signs. Language, therefore, depends upon a capacity to perceive the gesture as having a meaning, as being referential. Simple response to a gesture as a stimulus is not language.

The achievement of meaningful [iconic] tracing is the essential first step. The recognition that resemblance is produced by tracing allows the possibility of experiment with the act of tracing. By this practice traces can be made whose meaning is recognized by social convention. This is because in creating the trace of a bison another thing eventually comes into existence besides the trace itself, namely, the concept of depiction. The original trace makers must have recognized that what was going on was what we call depiction and that this was something new, although we have no basis for saying what they would have known it to be. The concept as a concept cannot be assumed to have been part of their knowledge.

A further matter eventually arises from the recognition of depiction as depiction, namely, the possibility of a meaning’s being given to something which is not depictive. We call this a code. An example of the process of encoding can be seen in the formation of neologisms in sign language, in which pantomime is rapidly reduced to convention in the establishment of a new term [see, e.g., de Matteo 1977:128]. Ucko (1988) has argued that there is no necessary sequence in the evolution of image-making style. We point out that once the practice of making coded images is established there can be no rules about the direction of stylistic change.

Depictive and non-depictive traces may exist in a system of shared meanings. For traces to have meaning they must be made rather than left. To make a trace implies attaching meaning to a mark—such as that made by a foot—which may or may not be the recognition that the mark implies the act of making it. The leaving of traces, in contrast, will normally be patterned—a footprint of one individual of a species is like the footprint of another—but, except among humans, there is no indication that the creature leaving the trace or any other individual creature attaches any meaning to it.

From the absence of mention in one survey of predation (Curio 1976), it seems that humans are the only animals to detect their prey through the marks left by that prey. One exception might seem to be the herring gull, a specialist predator which responds to “domes” that conceal crabs covered by sand (Curio 1976:98). It is not really an exception because the gull sees where the crab is through the sign, not where it has been. No crab, no dome.

The first signs that have been perceived as distinctively hominid, not belonging to the ancestor of an ape, are footprints in the tuff at Laetoli, 3.6 million years ago (Leakey 1981). This track consists, in part, of the traces of three bipedal individuals. One of these hominids planted its feet in the footprints of another. By this act, in which the feet seem to have been matched to the prints of the individual who passed first, we can infer that the one following perceived the earlier footprints in some way. We have no way of knowing how they were perceived or what, if anything, they were perceived as. It can be said that the first individual left signs of itself. It cannot be said that it set out to make those signs.

Once a footprint has meaning within a system of reflective language, any footprint can have that meaning, and so can a stencil of a foot or a drawing of a foot.

We cannot say what communication was like before depiction, and we can say little about it prior to the emergence of writing. We can, however, use the orders in which abilities emerge in the modern human infant as a guide, on the assumption that if a given modern characteristic B depends upon the existence of characteristic A for its emergence, it is reasonable to argue that A evolved before B as a general characteristic of all humans. It seems to us simpler to argue that, at some time during the evolution of modern humans, characteristic A might have developed into extinct characteristic C, although now it develops into B, than to argue that characteristic B once developed from extinct characteristic D, where now it develops from A. The early phase of the typical child’s life—around 18 months of age—in which drawing is not depictive is exactly characterized by context-dependent communication, including vocal communication. At the same time as young children begin to make marks recognizable as something—around 36 months—they begin to speak narratively and self-referentially [Ames et al. 1979], that is, reflectively. Typical three-year-olds seem to have what we call language. This does not imply a causal link between language and depiction in the development of modern children,
cause their language is acquired through their immersion in language-rich cultures. It is the developmental coincidence of depiction and reflective language that implies some connection between them.

The effect of depiction, we hypothesize, is to transform communication from a system of context-dependent signals, howsoever elaborate and organized as to syntactic sequence, into a system of reference independent of context—that is, a system productive of reflection and hence of narrative. In the development of children there is a tendency for the earliest utterances to be understandable in context and for later utterances to have meaning out of context but rarely to occur outside such context. The use of words in the absence of the immediate context of reference is a relatively late development in early childhood. The giving of meaning to the context-dependent utterances of very young children is a linguistic act by the hearer, not by the child (e.g., Lock 1980). These noises of very young children are not language. The ability of modern speaker-hearers to provide meaning for young children's utterances derives from such speaker-hearers' immersion in linguistic communities that not only have depiction but also, increasingly, have written language. This ability to provide meaning rests upon a history, and prehistory, of reflection. Depiction and its successor, writing, have been fundamental in the evolution of this ability.

The Social Construction of Mind

Recent treatment of depiction in the Upper Paleolithic (Stolar 1977, 1978; Halverson 1987:63–89) repeats the commonly expressed realization (see, e.g., Clark 1970:105–28) that this period represents a watershed in human evolution and that depiction plays a role in taking human life across some sort of threshold of reflectivity. Halverson (p. 68) quotes Cassirer, a philosopher who insisted on the centrality of language to the understanding of human life, as stating that “something enduring is determined and emphasized” in “the sign.” He himself remarks that the images in Paleolithic depiction “are abstracted from nature, yet concretely represent natural objects . . . [and] consciously created, . . . would invite a conscious response” (p. 70).

Sentiments like these we implicitly endorse, but we are producing a different argument about how depiction functions, namely, as the pivotal practice engendering reflective language. As against the view expressed by Halverson (and others) that “consciousness,” “thinking,” “imagination,” “memory,” and so forth, are categories of the mental, independent of language, we propose instead that these terms are grammatical categories (Ryle 1949, Wittgenstein 1958, Coulter 1979)—concepts which mark practices observable in linguistic communities. “Thinking” in this account is a concept which members of a culture use in the course of their everyday conversation to denote a form of conduct, the striking of a posture, or the statement of an opinion. The use of it does not mean that it is taken that the mind is an entity which does something called thinking, independent of such conduct. It is a way to talk about some of the things that people do and ways that they do them. Descartes's statement (1985 [1641]) “I think, therefore I am” is a line of rhetorical reasoning, that is, a statement in an argument about the question of his existence. It is not the statement of a discovery (the “discovery of the identity between thinking and existence”). Descartes and other theorists have denoted an element of the person as “the mental,” to distinguish “the mind” from “the body.” This is the expression of their belief in “the mind” as a real and proper object of study.

Olson (1986) traces this sort of move, in its turn, to the rise of literacy, especially of printed text, whose function in our sort of culture has been to generate the distinction between the given (the text) and the intended (the interpretation of the text). Olson has shown that preliterate human beings do not make that distinction, and the significance of that observation is to reveal the distinction's historical determination. It is only when the writing of text has come into existence historically that there can be a difference between the given and the intended in an utterance. It is not, therefore, a distinction universally applicable to any creature that happens to be biologically human, and this makes it improper (anachronistic) to assume “mind” (that which interprets the “given”) as a feature of humanness rather than as an element of post-literate linguistic practice. Hankoff (1980:9) sounds a similar warning with regard to the use by present-day theorists of “body-mind concepts” in discussion of the people of some 4,000 years ago.

Halverson's view takes these modern concepts for granted as being properties of the earliest image makers. This allows him to fancy the making of images as being the copying of a “remembered perception” (a mental image). (Dennell 1983:92) makes a similar statement. Furthermore, Halverson (1987:66–67) sees depiction as merely one of several means to the expansion of consciousness:

I am not suggesting that depictive activities were the cause of cognitive development toward conceptualization, certainly not the only cause, but they were surely contributing factors. I should imagine with Stolar (1977–78:25–26) that mimic activities and synecdochical responses to bones, for example, also contributed to the process. Indeed, any behavior including speech) in which subject and object, signifier and signified were consciously differentiated would have played a part. It happens that representational figures are the only surviving material indications of the process and hence of unique value in reconstructing it.

Our view is both more radical (about depiction) and more conservative (about “mind”). Depiction in particular provokes the reflectivity that in turn permits referential utterance. The significance of Halverson's closing sentence in the paragraph quoted lies in the fixity and endurance of these images for their producers and others. The images outlast, and hence persistently repre-
sent visible signs of, the gestures entailed in their making, as well as the things they are made and seen to resemble.

From the point made about “mind” as a social construct it follows that we do not accept the notion that such images are copies of a “mental picture.” Once there is an intention to depict, the repetition of depiction increases visually guided motor control. Satisfaction is gained each time the image formed resembles, and resembles increasing numbers of, recognized features of its object, within whatever criteria of adequacy are relevant to any occasion. It is subsequent to the production of this imagery and its effect upon the language habits of those making and seeing it that it becomes possible to address the notion of “imagined” forms at all.

Vygotsky [1978 [1930]] identified two types of memory in the developing child: that given naturally, in the form of mental imagery, and that given culturally through the social organization of mnemonic signs. He relied on Jaensch’s [1930] claims about the reality of “eidetic imagery” in children to bear the weight of his first assertion. Eidetic imagery has been taken to mean some form of “photographic” memory and is therefore, as Vygotsky himself says [p. 39], “very close to perception.” Subsequent study [see Gray and Gummerman 1975 for a review] has not borne out Jaensch’s claims about the nature, even the reality, of eidetic imagery as a unique phenomenon, and there seem to be no reliable reports of such memory imagery in children younger than six or seven [i.e., prior to their language’s being well developed]. This provides a basis for concluding that imagery is a property of children’s cognition only after they are competent language users and pictorialists and after they have been exposed to images, stories, and so forth. The apparent emergence of imagery subsequent to language and language-related development is consistent with our argument about how “imagined” forms evolve. Therefore we do not see a requirement to consider this sort of memory separately from the “social organization of mnemonic signs.”

According to Vygotsky [1978[1930]:39],

Even such comparatively simple operations as tying a knot or marking a stick as a reminder change the psychological structure of the memory process. They extend the operation of memory beyond the biological dimensions of the human nervous system and permit it to incorporate artificial or self-generated stimuli, which we call signs. This merger, unique to human beings, signifies an entirely new form of behavior.

Such signs are, of course, cultural objects for the simple reason that they have meaning and therefore are part and parcel of the communicative repertoire. They serve, as marks, to objectivate experience and “fix” it. The making of such signs has its own, reflective, effect:

Every elementary form of behavior presupposes a direct reaction to the task set before the organism (which can be expressed by the simple S-R formula). But the structure of sign operations requires an intermediate link between the stimulus and the response. This . . . is a second-order stimulus [sign] that is drawn into the operation where it fulfills a special function; it creates a new relation between S and R. The term “drawn into” indicates that an individual must be actively engaged in establishing such a link. This sign also possesses the important characteristic of reverse action (that is, it operates on the individual, not the environment).

The freezing or fixing of gesture [to use Leroi-Gourhan’s [1964–65] or Gibson’s [1966] sort of terminology], when that has meaning, represents the “invention” of memory. Reflective language, flowing from this practice, can thereafter take up part of the commemorative function of the fixed image. The use of the term “commemorative” is deliberate. Memorizing as a personal achievement follows from exposure to its public forms, in iconic, emblematic, or encoded displays, including the depictive products of other cultural members. This is, of course, an example of the general point Vygotsky makes, akin to Mead’s [1934], that “the mental” (in this case “the memorised”) begins in the social arena and, through rehearsal by members, becomes a more “personal” possession.

Olson [1986] notes the copious historical evidence, assembled by such scholars as Goody, Havelock, Innis, McLuhan, and Ong, for a correlation between literacy, particularly alphabetic script and particularly in mass printed form, and both religious reform and science. The reason for this [Olson and Astington 1986] is that the effect of literacy, in the public sphere, is to generate a distinction between texts and interpretations of texts. With the appearance of texts in lieu of speakers, meaning is not immediately available, and the interpretive act entailed in reading [as against the interrogatory freedom available in face-to-face conversation] becomes commonplace. Such practice provides the cognitive basis, in Olson’s view, for the description and interpretation of both scripture and nature, which in turn give rise to a perception of interpretation as a new category of activity, a category that points to the subject as separate from the object (the text). This in turn is a foundation for mind-body dualism, as we have already remarked.

Evidence in favour of this hypothesis comes from studies by Olson and colleagues of children’s responses to ambiguous stories. In one such story, Lucy [of Charles Schulz’s “Peanuts”), preparing to go to a party, says to Linus, “Bring me my red shoes.” Linus brings her a pair of old red running shoes, not her new red party shoes. Five-year-old children, just on the verge of reading, take it that what was said in such a story was what was intended, declaring that Linus did not do what Lucy wanted or what she instructed, which, they say, was, “Bring me my new red shoes.” Eight-year-olds tend to draw the same conclusion, though they make a more literal report of Lucy’s words. The great majority of ten-year-olds, however, can see that while Linus didn’t do what Lucy intended of him, what he did was what she asked. Olson goes on to say [1986:117] that “the child’s
conceptual distinction between the given (namely, the text or wording) and its possible interpretation . . . provides the basis for the child’s epistemological distinction” between what is seen and what is known. He here presents evidence that there is a positive correlation between children’s capacity to distinguish between what they see and what they know, in a test of visual ambiguity, and their capacity to distinguish between what is said and what is meant. It is the acquisition of this distinction that opens up the capacity to theorize about what is seen, not merely to perceive and talk about it. And it is public literacy, with its attendant demand for seeing the distinction between what is given and what is meant, that delivers this capacity. A significant aspect of Olson and Astington’s (1986) position and our analogous one is that post-literate concepts, like post-depictive ones, are cultural inventions and hence accessed by members whether or not, as individuals, they engage in the practice of writing or reading or depicting.

We have already argued that present-day theorists, taking this capacity for granted in their own conceptual apparatus, not witnessing its historical origin, assume “the mental” as an invariant design feature of Homo sapiens sapiens. But pre-literate humans do not see the distinction between what is said (or seen) and what is meant (or known), and from this we conclude that pre-depiction humans did not see the distinction between seeing a thing and responding to it. By “seeing” we mean what Wartofsky (1980:133) means when he observes that human vision is the product of the use of representations, which are visual artifacts (i.e., artifacts made expressly to be seen) . . . [and] it would appear that human vision is an artifact made by these other artifacts. . . . We create and transform the human visual system by means of the making of representational artifacts. . . . the human eye is a product of art. What may appear as a circularity here is, in fact, a feedback loop. The human visual system is thus both a condition and a product of human praxis.

Later he states that “human vision is not the neuro-physiological reflex of a causal chain starting with a light flux impinging on the eye and ending in the visual cortex. Though such a causal process may be a necessary condition of vision, vision itself is a human activity and not a neurological event” (p. 134).

The characterizing of human vision as a neurological event is the result of post-literate practices enabling theorizing about and investigation of the body. It is, once again, fallacious to take it that neurological forms of description provide an account at some more “primitive” level of existence.

Wartofsky’s argument is developed through reference to performance practices associated with the staging of theatrical scenes and the making of perspective depiction, but his general point (p. 134) aligns with our own: “Nonhuman animals cannot see what we see, insofar as our seeing has developed as a cultural skill, or art. What nonhuman animals cannot see is what we see by way of our picturing, and that is because animals do not make pictures.”

The Precursors of Language

Communication is common to many creatures, but only humans have the capacity to communicate their meanings independent of context. We suggest that this capacity derived from the fixing of gesture which showed a resemblance to the form of an object through mimicry in such a way that repeated reference could be made to that image and its resemblance to the object. This created memory and had an effect on the consciousness of pre-language hominids analogous to the effect of literacy on pre-literate consciousness. A similar sequence can be observed in the acquisition of language and depiction skills by children.

In the light of all this, the prehistory of tool making and use and of early image making and the relation of this record to the concept of language can be retraced as follows:

TOOLS

Many scholars emphasize the importance of tools in the emergence of language (Holloway 1969, Hewes 1973, Isaac 1976, Gowlett 1986), and we do not disagree that they were important. But it is possible, as Wynn (1980:788) has argued, to make objects which we see as tools without language. Indeed, Ingold (1986:349), in a passage which is otherwise not in tune with our argument, suggests that “we would . . . have to agree, with Pumphrey, that ‘the web of a garden spider and the nest of a chiffinch are highly fabricated implements quite as difficult to explain away as any product of lower Paleolithic man’ (1953:233). On the strength of the objects alone, we would have no more valid reason for assigning intellect to the human being than to the spider.” This argument would have been more appropriate if both writers had recognized (by avoiding the words “man” and “human being”) that we are dealing with the products of hominids, not modern humans. Even in modern humans, Ingold points out, much craft production follows procedures made routine by tradition and is no more taught through language than most language is.

Oakley (1954), in contrast, considered it impossible to “think effectively, to plan, or to invent, without the use of words or equivalent symbols,” and, since he distinguished between perceptual thought in chimpanzees (see a termite mound, get a stick to probe it) and conceptual thought in humans (or hominids) (set out to get food knowing that a stone tool would be useful at some time in the future), concluded that hominids would have had language, in contrast to chimpanzees.

Tool making and use in chimpanzees in the wild is now documented in a number of regions (McGrew, Tütin, and Baldwin 1979). In one wide-ranging comparison, 85–94% of tools were made from sources found within reach of the food source (about 2 m), although at another
site 98% of the tools came from more than 2 m. Besides this, Goodall [1968, cited by McGrew et al. 1979:200] observed tool selection 91 m from a termite mound and apparently [Goodall 1986:536] before sight of the mound. This seems like what Oakley would call conceptual as opposed to perceptual thought. But because chimpanzees do not have reflective language, the carrying of objects used for termite is not dependent on that ability, and this undermines Oakley's argument about hominids and language.

The longest recorded transport of a tool by a chimpanzee seems to be about 800 m [Goodall 1968; McGrew et al. 1979:200]. Hominids at Olduvai Gorge, about 1.8 million years ago, carried stone tools from the nearby volcanic uplands up to 10 km to the sites of DK and FLK N [Zinj] [Leakey 1971; Hay 1976:184], and some sites of this period had bones with stone tool cut marks but no stone tools [Potts and Shipman 1981] (the obvious inference being that the hominids were carrying stones with them). Hominids seem to have had some knowledge of what they were about to do. We do not think that they reflected on that. Their knowledge was procedural (knowing how), not declarative (knowing that) [Ryle 1949, Cohen and Squire 1980].

Binford's [1987] analysis of the Oldowan and Acheulean industries of East Africa suggests that from the earliest stages there were two patterns of tool manufacture and use: core tools like Oldowan choppers and Acheulean handaxes were manufactured away from many of the sites where they were found, implying that they were carried around as well as discarded; small tools [scrapers and cutters] were manufactured at sites where the core tools were also abandoned. This looks like a situation in which choppers and handaxes were tools needed quickly—say, in obtaining live animals which would flee by the time the hominids had seen them and then sat down to make their stone tools. The scrapers and cutters were manufactured at leisure, for example, after the lions had been chased away from the carcass of their prey. Language is not necessary to allow this behavior: natural selection would favor it.

With reference to the evidence from the Middle Pleistocene site of Zhoukoudian, Binford and Ho [1985:429] speak of "a noncultural form of adaptation which is strongly tool-assisted." Early hominid tool use was rather like chimpanzee tool use. All of the difficulties about the importance of tool making in the definition of the human species disappear if we look at it this way. Humans and hominids make and use (what only humans know as) "tools," and many other animals use "tools" [see, e.g., Beck 1980, Bonner 1980].

In addition to the conceptual problem of the distinction between hominid and chimpanzee tool making and use, there is a simple analytical problem. Holloway [1969, cited in Falk 1980a] suggested that early hominids were "imposing arbitrary form" on stone tools, and this statement became the basis for the further inference that tools "seem to indicate traditional transmission, productivity, duality, semanticity, and grammar" [Falk 1980a:75]. But it is the archaeologists who impose this arbitrary form, because the only way stone tools can be recognized is by recognizing pattern. There is therefore a great emphasis on those artefacts in Oldowian assemblages which appear to have an imposed form, especially the pebble choppers. It is at least arguable that the technology of flake production, rather than any desire to produce objects to a predetermined shape, imposed some of that form. Dibble [1987] has shown that several of the typological categories used by archaeologists in classifying Mousterian stone industries actually result from interruptions to a single process of reduction, during use and resharpening, of a more limited number of forms. He concludes that the typologies are measures of intensity of use rather than reflecting the intended production of artefacts with particular shapes. This raises serious doubts about the existence of "mental templates" for tool production in the Middle Paleolithic.

In the face of such doubts, it seems unlikely that we need to invoke "mental templates" for the handaxes of the Lower Paleolithic. On the other hand, no one has yet raised such doubts about the tools of the Upper Paleolithic: the novel bone tools, the new technique for producing large numbers of relatively uniform flaked-stone blades, and the new range of retouched stone blades. If there is a case for the use of language in the manufacture of tools, it is to be found in the Upper Paleolithic and not before.

We may therefore conclude that the main importance of early tools for our argument is not in their non-utilitarian status [Edwards 1978] but in the sequencing of actions entailed in their manufacture.

**Sequencing**

There are two main positions about the timing of the origins of language. One suggests that early hominids had language [Falk 1980a, b, 1983, 1985, 1987a, b; Tobias 1981]. The other, that argued here, suggests that language is a late phenomenon in hominid evolution, probably as late as the Upper Pleistocene. The first argument depends on language as a major cause of brain expansion; the second suggests that language must have followed brain expansion.

Falk [1983] showed that KMN-ER 1470, dated to before 1.8 million years, was the earliest known hominid with a cortical sulcal pattern similar to that of humans in the region of Broca's area. This area is associated with speech in modern humans, but not only with speech. According to Passingham [1981], the "primary function of Broca's area is to direct the sequencing of the basic movements, in this case the sequencing of the vocal cords." It is implied that the emergence of this area in the brain was a prerequisite to the ordered pattern of utterance that now characterizes human speech. As we have shown, more is involved in sequenced communication's becoming language.

Gowlett [1984a, b] documented the extensive chains of action necessary to make a handaxe. There seems little doubt that control of the complex sequences of operations required in producing stone tools was an
tential part of the preadaptation for the more complex sequencing control necessary in production of speech, the associate of language. Making stone tools played a key role in the evolution of brain structures controlling other sequenced behavior [Bradshaw and Nettle-ton 1982, MacNeilage, Studdert-Kennedy, and Lindblom 1984].

Calvin (1982) has argued that aimed throwing provides a context for natural selection to favour both handedness and multiplication of the “numbers of redundant timing circuit neurons in sequencer cortex” [Calvin 1982:118] which would produce precisely the expansion which Falk observed in the region of Broca’s area. Expansion of this region of the brain would also select for skilled application of blows in stone knapping, fine control of hand movements for gesture, and the control of oral-facial musculature essential to the production of speech. All of these advantages in foraging, technology, and communication can occur without turning communication into language, but all are essential to the creation of depiction through mimicry which did.

Wynn (1979, 1981) has identified, in the changes in stone tools between the Oldowan and the Acheulean, stages which Piaget described in the development of intelligence in modern children. Even if we accept Wynn’s analysis, we still need not accept that these stages existed in the hominids. The fact that we can recognize, say, that an object is symmetrical does not mean that the makers knew it was symmetry they were producing.

EARLY IMAGES

In the recent development of ideas about the significance of European Upper Paleolithic art, several archaeologists (Conkey 1978, 1980, 1983, 1985; Gamble 1980, 1982, 1983, 1986; Jochim 1983; Gilman 1984; Davidson 1989) have worked on the idea that the knowledge encoded in the “art” was not intrinsic to the object but part of a system of communication which included such objects. The existence of rules can be seen in the Venus figurines of about 25,000–23,000 years ago. Gamble (1982, 1986) has drawn attention to the significance of the common features of these figurines: they lack faces and feet, and there are no males. The distribution of these conventions across a region some 500 km wide and stretching 3,000 km from the Atlantic Ocean to the River Don requires a shared understanding across that region. We cannot detect any sort of convention, with or without exceptions, earlier than the Upper Paleolithic.

The earliest group of objects for which a convention can be detected through the presence of repeated patterning is the scratched bones from sites contemporary with the three-dimensional figurines of Vogelherd (Hahn 1972, 1983), and the so-called vulvas of the Aurignacian in Dordogne [Delluc and Delluc 1978, 1981; see Bahn 1986]. Clegg (1986), in a comment on Bahn’s [1986] paper, says that the signs in the great panels at Cathedral Cave, in central Queensland, which consist of two concentric ovals, were called “macrozamia nuts” [Goddard 1941], whereas he was inclined to call them “vulvas.” It does not matter, however, what was intended by the artists. The interpretation of these signs, as of the Aurignacian ones, is not, and therefore was not, intrinsic to the objects themselves. Repetition of the pattern of an oval with a groove or score mark at the seven Aurignacian sites indicates a convention shared within a community, whether or not there was any intention to depict a vulva. The referential opacity of the signs shows that this depiction was already the subject of encoding processes and therefore cannot be early in the history of depiction. But the Vogelherd figurines cannot be the earliest depictions either, since they required well-developed depictive skills. The repetition of scratched marks on bones at 25 sites including Vogelherd is evidence of intention to make a pattern that the scattered examples from the preceding 60,000 years are not. This repetition may also be the result of encoding processes, which are necessarily later than the earliest depiction. Our hypothesis is that images with referential transparency exist earlier than the vulvas and the figurines.

The argument we have presented about the role of depiction in the origins of language suggests that there was a fundamental importance in the production of marks which resembled things in the natural world through the freezing of gesture to make a trace, perhaps initially by accident. Only after this occurrence had led to the deliberate making of an image could meaning be given to non-depictive marks, and only after this could the concepts of depiction and abstraction exist. The transformation of the capacity to make deliberate marks into the capacity to make them symbolic seems to have taken place in the first half of the Upper Pleistocene. We have shown that this was the period in which hominids became human through the emergence of reflective language.

Conclusion

Speculative hypotheses always run the risk of being impossible to prove. We do not believe that this argument will stand or fall in response to any new data which may be found in the future. We have made one prediction, that depictions earlier than those dated around 32,000 years should be found and that these should be able to have been made as a result of mimicry. Failure to find such images, of course, will not constitute falsification.

Rather, we suggest that the virtue of our argument is in its power to explain the associations between features of the archaeological record. In a recent interview, Binford [Renfrew 1987] pointed to the complex of organizational changes marked in the transition between the Middle and Upper Paleolithic and suggested that they
require an explanation in terms of the emergence of language. He recognized the need for consultation between archaeology and psychology in constructing such an explanation, and that is what we were already seeking to do.

The power of the argument is in its ability to account for the emergence of depiction at the same time as the emergence of many other features. We would go further and suggest that the late emergence of reflective language accounts for the nature of the break in the rate of change in artefact technology, ritual and burial [Harrold 1980], regional variation in style [Gamble 1986:322–42], social differentiation [Davidson 1989], the emergence of hunting and gathering [Foley 1986] and of agricultural systems [Davidson n.d.], and the migratory capacity of humans in colonizing Australia [White and O'Connell 1982, Thiel 1987] and the New World. Far from agreeing with ethologists and others that some animals have culture [e.g., Bonner 1980], we suggest that there can be no such thing as culture without language and the socially determined sharing of meaning and value. It will therefore be misleading to talk of culture for any hominids before fully modern humans.

Paleoanthropologists discuss the origins of fully modern humans in terms of the evidence of skeletal structure [e.g., Valladas et al. 1988, Stringer 1988, Stringer and Andrews 1988] but with increasing emphasis on behaviour [e.g., Trinkaus 1986]. Lieberman [1984] has shown how the anatomical changes implied in the evolution of fully modern human skeletal structure would have permitted a wider range of vocal sounds. We agree with Rindos [personal communication, June 10, 1988] that the final step involved changes which were not fundamentally genetic. However, acquisition of language would have provided a selective context which favoured those hominids which had more effective communication. We have no doubt that, once acquired, language could have been more easily taught [with or without depiction] to, for example, Neandertals than it is today to chimpanzees. But the limited vocal range of Neandertals [Lieberman 1984] would have selected against their genetically determined physical characteristics very rapidly. Our argument has been that the emergence of language, through depiction, provided the behavioral context for selection to favour the greater vocal range.

Comments

DAVID F. ARMSTRONG

Gallaudet University, Washington, D.C. 20002, U.S.A.

10 X 88

Davidson and Noble present an appealing theory on the relationship between the origins of depiction and the origins of human languages and implicate the beginning of the Upper Paleolithic as a crucial period during which “reflective” language first emerged. The critical element in this speculative theory is the “freezing” of iconic gesture through depiction. Physical evidence supportive of the hypothesized chain of events includes the appearance at this time of the first preserved examples of recognizable representational art and changes in the human upper respiratory tract believed to be indicative of increased articulatory efficiency in speech.

Others have pointed to the beginning of the Upper Paleolithic as a likely time for the origin of languages recognizably like those of modern humans. For example, Hockett [1978], employing some of the evidence adduced here, arrives at roughly the same conclusion. Davidson and Noble note that their theory puts them in the “punctuationalist” camp as opposed to the “gradualist,” and this contrast reveals some problems with their formulation. The fundamental issue is whether language developed gradually through a series of stages over a very long period of time [e.g., Armstrong 1983] or arose rapidly relatively recently. The first of these theoretical frameworks assumes that language-like abilities were present in the common ancestor of humans and African apes; the second assumes a discontinuity, language being a unique attribute of human beings. This latter also calls for a triggering mechanism, genetic or behavioral. Davidson and Noble recognize that a genetic trigger will be almost impossible to identify. As they point out, language is not visible in brain structure, but cortical elaboration in the region of Broca’s area is, and it is visible well before the Upper Paleolithic. Similarly, with respect to the upper-respiratory-tract evidence, it can be argued that the changes leading to the modern human condition were gradual and started well before the crucial time [e.g., Laitman 1983].

That leaves us with a need for a behavioral trigger, and Davidson and Noble present a plausible argument for the discovery of depiction as the key to the quantum leap from some precursor to fully “reflective” language. The problems are several. First, they are vague with respect to precisely what that precursor would have been like, although they suggest some form of mimetic gestural system. Second, their argument hinges on universal (or near universal) distribution of depiction (as they define it) in nonperishable media among all human societies known to have had fully developed languages (i.e., all human societies extant within, say, the last 10,000 years). This is required because, according to their argument, depiction and speech must scale. If speech is present, depiction must also be present; and the depiction must be in nonperishable media, because the possible presence of depiction in perishable media (but not simultaneously in nonperishable media) prior to the Upper Paleolithic is dismissed. They have not presented evidence that would support this requirement. Third, the argument as presented does not conclusively solve a fundamental problem faced by all such formulations: “Human language is an embarrassment for evolutionary theory because it is vastly more powerful than one can account for in terms of selective fitness” [Premack 1986:133].

This third problem is potentially the most damaging.
for an evolutionary argument that is based on a single, relatively simple triggering mechanism. Central to the argument of Davidson and Noble is the idea that language-as-we-know-it could not have existed for long before the Upper Paleolithic—otherwise, the elaboration of material culture, the development of ritual, etc., would have started much earlier. Implicit in Premack's statement is the observation that the linguistic capacity to support the technology of the 20th century existed but was not revealed in the material culture for many thousands of years. Therefore, the absence of virtually any sort of material evidence, including depiction, cannot be taken as conclusive evidence for the absence of fully developed languages. Conversely, Davidson and Noble correctly recognize that the simple technology of the Lower Paleolithic in no way required language to support its development.

The selective advantage accruing to users of human languages must lie elsewhere than in the material culture, precisely in the abilities to develop and elaborate complex social structures, plans, and strategies that belong to all those who live in human social groups. It is difficult to conceive that these things do not have a very long evolutionary history. Davidson and Noble have advanced the theoretical development in this area by citing those such as G. H. Mead who recognize the role of social life in the development of language, but their argument that depiction is the material residue of the initiation of the process is not entirely convincing. They have, however, taken another important step in the process of identifying the critical factors in the origin of language, supporting the argument that language must certainly have been in existence by the beginning of the Upper Paleolithic, and sharpening the debate between those who believe its development must have been a gradual process and those who believe it was punctuated.

L. T. BLACK
Department of Anthropology, University of Alaska,
Fairbanks, Alaska 99775-0160, U.S.A. 20 x 88

The paper under discussion caused me to experience an acute nostalgia for the good old days when the French Academy would accept no papers on the origin of language because no evidence could, by definition, be forthcoming and when American and British anthropologists zeroed in on synchronic interpretation of cultures. Even the latter appeared to me for a moment more desirable than the jargon-laden speculative outpourings produced by Davidson and Noble.

The activity of the human brain associated with complex behaviors such as depiction and use of language (I eschew the use of the nonsensical term “reflective language”), not to speak of the human capacity to operate with concepts, is not clearly understood even by the leading specialists in neurophysiology today. I urge that persons who wish to address the emergence of such complex behaviors on the part of members of Homo sapiens or the genus Homo acquaint themselves with the findings and interpretations of the leading neurophysiologists prior to proposing grandiose interpretive schemes. The works of Eric Lenneberg should be a good start. I think that in view of these findings, the simplistic sequencing of development of structures prior to development of functions and depiction prior to linguistic communication will be found wanting.

All of us who read works devoted to the archaeology of the Upper Pleistocene know that a very long period of relative uniformity in tool manufacture is supplanted by spatial and temporal differentiation. We also know that this period is marked by the production of a variety of artistic forms. We have no good explanation for these developments. The speculative suggestion is made that they are best explained by the acquisition by H. sapiens of a linguistic communication system triggered by depictive activity that originated in mimicry of animals. How the Cro-Magnon artist made the conceptual transition from hunching his body like a bison to the idea that he could reproduce the bison by drawing is not made very clear to me. This suggestion, therefore, will remain just that: unfounded speculation.

I am saddened that there seems to be developing an academic network of persons dedicated to the propagation of the idea that language is a late evolutionary development linked to the emergence of prehistoric art in Europe. Though this propaganda cannot be substantiated by evidence, the space devoted to it in various scholarly journals and its dressing up in pseudoscholarly trappings will only lead the unwary beginning scholar into unproductive bypaths. Let us return to the sound approach of our forebears and not waste time on unresolvable chicken-and-egg questions.

WILLIAM H. CALVIN
Biology Program NJ-15, University of Washington,
Seattle, Wash. 98195, U.S.A. 19 x 88

I agree that depiction is a key step and that it likely emerged from a primate base of neural sequencing machinery [as suggested by the monkey’s left-hemispheric specialization for temporal ordering [see review by Falk 1987]], gradually augmented over the last few million years by natural selection for hammering and throwing abilities (Calvin 1983, 1986). If depiction is a secondary use of neural sequencing machinery primarily under natural selection via ballistic skills and not itself contributing much to physical adaptations, then simulating referential-depiction-actor-action-agent relationships inside our heads could have arisen as late as the Upper Paleolithic.

Yet depiction could have been quite early. I suggest that animal tracking was a likely starting point, given our impoverished sense of smell, our ancestors probably relied more on visual clues than on the olfactory ones used by carnivores. They looked for hoofprints, droppings, hair snagged by brush, dust newly settled on
leaves, such may not "represent" the shapes of the animals, but carnivores [and surely our hominid ancestors] had no difficulty in associating the tracking sign with the real thing. Our ancestors probably noted tracks of fellow hominids as well as of prey animals, maintaining an awareness of other hunting bands. And returning hunters or gatherers trying to rejoin a band on the move would have needed to interpret the trail left by the main group in order to catch up.

It seems a small step from regularly noting such a hominid track to modifying one's own track, e.g., dragging or stamping the feet on one branch of a trail junction so as to aid a trailing member attempting to rejoin the band. Such protodepiction constitutes a stand-alone communication, to be interpreted at some later time, about what someone should do. An elaboration of it might have occurred in the context of a hunting-and-gathering band on the move, in which two members who had seen different promising food sources earlier in the day were trying to persuade the group to take different branches of a trail. Mimicking the signs left by the animals seen, as by each scout's trampling few to many footprints in the dust leading in the direction he favored, would have been a simple way [analogous to the bee's dance [von Frisch 1967]] of informing the others about the potential richness of the food sources. The group could then have chosen the more promising branch on the basis of quantity [or perhaps type] of mimicked footprints. All this could have been done without concepts of number or use of arbitrary symbols—but could have advanced toward them via a series of intermediate stages, in much the way that the modern corporate logo has drifted away from a realistic product depiction.

Elaborating such messages was likely aided by some neural sequencing machinery, essential for ballistic movements but borrowable for other sequencing uses. What this "serial buffer" ordinarily chains is elements from its repertoire of "motor tapes"—usually with output pathways inhibited [as in our nocturnal dreams] but with a "good-enough" command sequence [Calvin 1987, n.d.] occasionally let loose to create a real series of linked movements. To me, the interesting step is juxtaposing sensory schemas with the motor schemas to get more general actor-action-object sequences. An elementary one is the clause, a subject-predicate sequence in which the predicate is usually an action [a "verb"] plus an object of the action [or sometimes an adjective denoting state of being]. Originally, the actor-subject would always have been implied: oneself. But sensory schemas ["nouns"] would often have been associated with the object of an action ["Run after that rabbit," "Hammer that nut," "Throw at that gazelle"], suggesting that the primitive sequence is simply the predicate part of the clause: the verb and its object. Using nonsensory schemas in a subject role seems the important step because it allowed mental simulation of the actions of others. Yet it may not have been such a leap because of implied subjects: dragging one's feet in the dust at a trail junction is essentially saying, "You turn this way," or "The animals went that way," depending on context.

Acquiring an external stand-in for an internal schema could eventually have extended to familiar elements of both the sensory world and the movement repertoire. While natural selection could have promoted this sign-for-schemas vocabulary building, so could sexual selection; a pair of chimpanzees sometimes establishes an away-from-the-group "consortship" (Goodall 1986:453) near the onset of estrus. One can imagine that "Let's go that way when Alpha isn't looking" might have been communicated by a stare, eye movements, and postures [especially facial expressions]. Vocalizations seem particularly likely for primitive "adjectives," intensifiers, and negations, as species-typical exclamations often express anxious-angry-happy-hungry and "Don't do that." [Goodall 1986:127]. More advanced media of expression would be gestures [cf. Hewes 1976] and depictions in the dust, with innovative vocalizations to represent nouns and verbs perhaps coming later. It would have been a "mixed-media" vocabulary: a scout arguing at the trail branch might have used a stare to denote the clause's subject ["you people"], a body posture as verb ["go"], and a protodepiction ["toward the animals"] as object, perhaps with a vocalization ["food cry"] as intensifier. Juxtaposing sensory, motor, and state-of-being elements seems to carry the clauselike message here—mere association, not the ordering [syntax is not one of the conspicuous features of modern sign language, either].

When did the ordering of the elements begin to carry an additional implication? Perhaps it came with specifying the subject: in a clause with two nouns ["John called Julie"], which is the subject, which the object of the verb? A nonimplied subject often gets one into ambiguity unless there are some local customs governing which noun is the actor and which is the acted-upon. Though inflections of voice or gesture can denote "object," all six permutations of English's subject-verb-object syntax are in use somewhere [Smith and Wilson 1979:204]. If a noun could be expressed in more than one medium, the one utilized could also specify noun-as-subject or -object. It is prevention of ambiguity between subject and object ["Julie called John" has an entirely different meaning], not syntax as such, that greatly expands the number of possible messages that can be transmitted and understood.

Duality of patterning [phoneme order to make the less frequently used words, word order to make novel sentences] tends to suggest that phoneme sequencing is more primitive. But it seems possible from such protodepiction examples that grammar could have arisen without elaborate vocalizations, that the malleable message could have preceded the modern medium. Thinking our way through such possibilities may be more useful than defining thresholds for True Language and True Culture or focusing on particular routes of expression such as depiction, gesture, eye-facial-body posture, and vocalization.
As the resemblance rate, move, what requires notion. Davis hominid judgment “nonnatural” notion. Evanston, cognitive Department WHITNEY I40 appreciating even “trace” logic. Davis 1986a. The cube appreciation it. The cube square origins and making a best guess. It is open to my 1986 argument to see the perception of resemblance as mediated linguistically. At any rate, I fail to see that Davidson and Noble escape the very criticism that they aim at me. The perception of resemblance operates at several levels in their scenario, and supposedly prereflectively.

As an analytic strategy, I adopted my primitives (e.g., seeing-as, disambiguation) as far better than another one that Davidson and Noble must build into their account. As does another recent pseudosolution to the problem of the origins of depiction (Lewis-Williams and Dowson 1988:213–16, 236–37, and see Davis 1988b), theirs must assume the hominid’s ability and intention to refer. By begging the key question about reference, it is almost totally empty as an explanation; although it may not run afoul of the psychology of perception, it is laid low by the logic of depiction.

The surface of the moon resembles a human face but does not depict it. An Egyptian hieroglyph with the consonantal value m resembles an owl but does not depict (or even otherwise denote) it. My signature resembles—is a “trace” of—the gesture that produced it but does not depict it. A film I make resembles a dream I had but does not depict it (describing rather, say, the objects my dream represented). A square drawn on a sheet of paper resembles both a cube (seen front-on) and a pyramid (seen base-up) but depicts only one or perhaps neither of these possible solids. And so on. If the graphic pattern (visual display) depicts the object, then its relation of resemblance to that object can be governed by straightforward rules of projective geometry [Hagen 1986], but resemblance alone, however achieved, is not enough to give depiction. Roughly speaking, if and only if the resemblances in turn stand for that which they resemble (and usually only one of the objects that could be so described) do they depict. A great deal can be said about the distinctive logical properties of depictive vis-à-vis other forms of denotation (e.g., about their “density” and “repleteness” [Goodman 1970; Bach 1970; Elgin 1983: 141–83]). In this light, in Davidson and Noble’s scenario for the origins of depiction the following must hold:

1. Gestural mimic resembles the object but also stands for it via those resemblances; otherwise it could denote any one of the other things that it also resembles (e.g., a hilly horizon rather than a bison’s back, the same gesture as was made yesterday). The gestures must be tied somehow to objects ostensibly or indexically. For this crucial feature of the “mimic” (mere copy), they use the term “indication,” giving their whole game away. The need for resemblance in this “indication” is not explained. Why not just point to the object? Why indicate by waving the hands in a curve describing the bison in such rich modulation, with such density and repleteness of morphology [Goodman 1970]? For the very first stage, supposedly predeceptive and prereflective, evidently the gestures required are already depictive! Their “mimics” provide information to a viewer about the size, shape, texture, etc., of an object (rather than merely deictic data), albeit in balletic rather than graphic form—a “signalling” by pictures.

2. The “trace” of the gesture in a permanent medium resembles the object but also stands for it via those resemblances; otherwise it could denote any one of the other things it also resembles (e.g., the gesture producing it, which it does not depict), denote without resemblance (e.g., serving as signature denoting someone’s presence), or not denote at all (e.g., serving as ornament).

3. The “signs” repeating, varying, and conventionalizing the “trace” resemble the object but also stand for it via those resemblances, for the same reasons as above.

1. A sophisticated mimetic theory of depiction appears in the writings of Gombrich (e.g., 1961, 1973). Goodman’s objections (e.g., 1970) are devastating. Peirce’s concept of an “iconic sign”—at least if this sign is regarded as attached to its object by resemblance only—must also be regarded as incoherent (Peirce 1932, Amyx 1955, Morgan 1955, with criticisms by Burks 1949, Bierman 1962, and see Ayer 1968:140). Of course, similarity between sign morphology and stimulus properties is still one necessary if not sufficient condition of depiction (see Knowlton 1966, Manns 1971, Schwartz 1980). Tversky and Gati (1977, 1978) and Rosch and Lloyd (1978) show that similarity judgments are influenced by the task, the directionality of the judgment (Stimulus A to B, B to A?) and the “pre-existing classification context.” Such factors can be disentangled empirically. Some of them, of course, may be linguistically mediated—a problem for the authors’ but not for my account.

2. It is unclear whether their “signalling” involves reference—whether they are invoking Peirce’s notion of an “indexical sign” or some other species of reference (see, in general, Evans 1979, Devitt 1982).
“In creating the trace of a bison,” Davidson and Noble say, “another thing eventually comes into existence besides the trace itself, namely, the concept of depiction.” More generally, they view reflexive language as deriving from “the fixing of a gesture which showed a resemblance to the form of an object through mimicry in such a way that repeated reference could be made to that image and its resemblance to the object.” Quite. But whence this “reference”? It is obvious how mere morphological resemblances can be preserved from object to gestural mimic to “trace” to “sign” through the causalities of the making processes as such. But as far as I can tell, the authors provide no account of how the manufactured resemblances do or come to denote the original object. They either assume reference (by smuggling in “indication” or pictorial mimics and traces) from the beginning, create it suddenly and without analysis (a nonsignifying “trace” becomes “sign”?), or simply beg the question altogether.

Furthermore, I see no necessity to their sequence (gestures to traces to signs). If a gestural mimic refers via resemblance to an object, then it is an arm-waving picture; if a trace refers via resemblance to an object, then it is an imprint picture; if a sign refers via resemblance to an object, then it is a drawing picture. The logical properties of depiction are quite independent of its physical medium. It is therefore cumbersome, uneconomical, and implausible to require that gesture precede trace precede sign, especially if, as I suggest, all must already be depictive in their account.

I agree with Davidson and Noble that depiction, once inaugurated, will have profound consequences for individual cognition and for social life. Unfortunately, they assert that it requires a shared system of meanings. Someone can make and see a depiction that no one else can interpret as such, that is never communicated to anybody; the logical properties of a picture are independent of their communicability [see Searle 1986]. Moreover, sheer repetition of the morphology is not sufficient to establish “shared” or “constant” meaning [recall the square as “cube” for you, “pyramid” for me]. That one maker can pass on the meaning of a picture to another—for the “chain of designations” [Devitt 1982] to have a social life—should not be assumed as intrinsic to depiction. Facing the well-known paradoxes of mutual knowledge and the emergence of meaning conventions [see Biro 1979, Smith 1982, Sperber and Wilson 1986], a further analysis of a context of communication and interpretation must be carried out to explicate the sociology or “epidemiology” of representations [see Sperber 1985 and, for evolutionary essays, Grice 1982 and Tennant 1984]. Throughout Davidson and Noble’s presentation I sense conflations of maker’s meaning with “meaning in a language” [see Grice 1968, Schiffer 1972], of repetition with convention [see Lewis 1967], of meaning with communication [see Searle 1986]—to the extent that an important ingredient of Upper Paleolithic society, the “sharedness” of its meanings, may just materialize out of nowhere, like reference itself.

Although my conclusions differ from Davidson and Noble’s, I have very much enjoyed reading their interesting contribution. Unfortunately, what it is going to take to settle the debate about when language originated in hominids is a time machine. Until one becomes available, we can only speculate about this fascinating and important question.

The suggestion that depiction in the archaeological record indicates the existence of language is convincing. However, this does not mean that lack of depiction during the Lower Paleolithic can be construed to imply lack of language.

Since human language depends on a lateralized brain, it is instructive to investigate cortical lateralization by comparing brains/lateralization in living primates [including humans] and by study of the hominid fossil record [Falk 1987b]. Although those of us who study hominid brain evolution (“paleoneurologists”) are notorious for our disagreements [Falk 1987a], we do seem to agree that early hominids may have been capable of language [Tobias 1981, 1988, Holloway 1983a; Falk 1983, 1987a, b]. What accounts for this unusual agreement is the strength and convergence of biological/comparative evidence that favors an early origin for language.

Cortical lateralization is not new with hominids. It appears to have had a long evolutionary history within primates. MacNeilage, Studdert-Kennedy, and Lindblom [1987] surveyed studies from prosimians to apes and found that left-hand preferences, when they occurred, were for visually guided reaching for food that “may reflect a spatio-motor specialization of the right hemisphere related to the right hemisphere visuo-spatial specialization of humans.” They found that right-hand preferences occurred for some monkeys and apes but not prosimians and were for acts that required manipulation and/or practiced performance. Skeletal asymmetries have recently been detected in the forelimbs of rhesus monkeys and are in keeping with the notion that this species may rely preferentially on the right hand for certain manipulations [Falk et al. 1988]. These findings are also consistent with reports that split-brain rhesus monkeys show right-hemisphere superiority for differentiating photographs of monkeys’ faces [Hamilton and Vermiere 1985]. Other studies of Macaca report left-hemisphere dominance for processing species-specific vocalizations [Petersen et al. 1978, Heffner and Heffner 1984]. Thus, at least one genus of Old World monkey exhibits left-hemisphere dominance for communication and fine motor skills and right-hemisphere dominance for visual processing. This implies that the neurological substrate for the basic verbal/visual dichotomy that characterizes Homo sapiens may indeed have been present during early anthropoid evolution [MacNeilage et al. 1987, Steklis 1985, Falk 1980a].

But monkeys don’t have language and humans do. Are
there morphological manifestations of human brains that (a) correlate with functional lateralizations including language and (b) are capable of leaving traces in the hominid fossil record. Indeed there are. Shape asymmetries of the frontal and occipital lobes, known as petalas, exist in human brains (and to a lesser degree in brains of monkeys and apes) and are statistically associated with handedness in humans (LeMay 1977, Calaburda et al. 1978, LeMay, Billig, and Geschwind 1982). Further, a characteristic sulcal pattern associated with Broca’s speech area in left frontal lobes is present in human but not in ape brains (Connolly 1950, Falk 1983). Both humanlike petalis (Holloway and De La Coste-Lareymondie 1982, LeMay 1976, Tobias 1988) and the pattern of sulci associated with Broca’s area (Tobias 1981; Falk 1983, 1987a) have been detected on endocranial casts (endocasts) from the early part of the hominid fossil record.

The oldest evidence for Broca’s area to date is from KNM-ER 1470, a H. habilis specimen from Kenya, dated at approximately two million years ago (Tobias 1981, 1988; Falk 1983, 1987a). From that date forward, brain size “took off,” i.e., increased autocatalytically [Godfrey and Jacobs 1981] so that it nearly doubled in the genus Homo, reaching its maximum in Neanderthals (Falk 1987a). If hominids weren’t using and refining language I would like to know what they were doing with their autocatalytically increasing brains (getting ready to draw pictures somehow doesn’t seem like enough). Davidson and Noble seem reluctant to accept Oldowean assemblages as evidence that early hominids imposed arbitrary form on stone tools (Holloway 1969) and suggest instead that the perceived form is either in the minds of the archaeologists or a by-product of flake production. Acknowledging that hominids carried stone tools with them and therefore “seem to have had some knowledge of what they were about to do,” they “do not think that they reflected on that.” I think that they probably did and want to be first in line to travel on the time machine.¹

MARY LEcron FOSTER
Department of Anthropology, University of California, Berkeley, Calif. 94720. U.S.A. 11 X 88

Davidson and Noble are right that a faculty for mimesis is crucial to the genesis of language, but I find their argument that language and culture evolved from pictorial iconicity untenable. They argue that communication is derived from the “fixing of gesture which showed a resemblance to the form of an object” ([italics mine]. This illustrates a commonplace but mistaken assumption that language is essentially object naming and must have begun as such.

Upper Paleolithic art is not depiction in the traditional sense of the term. Parts of animals rather than wholes are represented, distortions are conventionalized, and abstract signs abound, suggesting figurative rather than referential intent. Davidson and Noble argue that distortion and lack of parts in early “Venuses” suggest symbolic coding, but this is a far cry from the depictive mimesis that their theory requires. Realism is deliberately sacrificed, for undetermined reasons, suggesting a sophisticated symbolism that had developed far beyond its early beginnings. The fact that the earliest known art begins with this opacity and progresses toward greater realism undermines the depictive argument. It must therefore be assumed that language also began with greater abstraction (i.e., opacity) and moved toward specificity.

Language, as Davidson and Noble emphasize, is a sequential activity. The most crucial aspect of language structure—and the most overlooked in debates about whether apes are capable of language—is the systematic organization of morphemes and words into classes (paradigms) on the basis of processual similarities. Generational rules for language use are based on the fact that there are classes, although generational grammarians tend to ignore the systems underlying the classes. The ability to distinguish and utilize class similarity is the crux of language, as pregnerational linguists such as Sapir and Whorf recognized, and is the product of a capacity to exploit mimesis in abstract ways.

Despite the dangers of an overstrict recapitulationist interpretation of ontogenetic stages, ontogenetic data can provide significant clues to language evolution. Present-day caregivers furnish modern children with context-appropriate, ready-made words, largely nouns, but psycholinguists find that children typically overextend their meanings. First nouns are applied to situations perceived as having similar characteristics. Vygotsky (1962:73–74) exemplified this focus on abstract properties with a child’s use of sutki [meaning “day and night” in Russian] to mean first a seam in cloth, then any junction, such as a corner, then twilight (the joining of day and night), and finally, accurately, the time from one twilight to the next.

This ability to abstract spatiotemporal characteristics from specifics and to generate these to other areas is what gives language its generative productivity. By 12 to 14 months, roughly the time of first overextensions, infants are able to substitute one object for another in play-miming of adult use (Shotwell, Wolf, and Gardner 1980). Syntactic substitution of an element with some abstract similarity to the other is a crucial characteristic of language.

While most ape studies exhibit the same fallacy as is found in the arguments of these authors—an implicit assumption that object naming constitutes language—they also demonstrate that apes possess a mimetic ability to substitute one object for another and to recognize similarities in the use of different objects. These are prerequisites for language.

Language reconstruction demonstrates that differential manipulation of the oral tract originally conveyed meaning, with individual sounds roughly replicating movements and spatial relations in nature (e.g., Foster

¹. Supported by Public Health Service grant 7 RO1 NS24904.
PAUL GRAVES  
Department of Archaeology, University of Southampton, Southampton SO9 5NH, England.  
30 IX 88

Whilst I find much to agree with in the methodology adopted by Davidson and Noble, I cannot accept their two central arguments: that there is a unitary “origin” for “modern” man and that this origin can be identified with the appearance of depiction in the Upper Paleolithic.

To begin positively, I endorse the emphasis on greater interdisciplinary cooperation between archaeology, psychology, and anthropology. I believe that this is essential for a full understanding of human evolution. Davidson and Noble’s critique of what one might call cognitive uniformitarianism is a case in point. The assumption of homologous mental processes in early and modern hominids has seriously handicapped attempts to explain how these processes evolved. Thus, as the authors state, interpretations of lithic technologies often make an unnecessary appeal to the arbitrariness of geometric form (e.g. Gwollett 1984b). I suggest that such forms are an excellent example of Gibson’s theory in practice; they are the product of a transaction between the tool maker and the “affordance” (sensu Gibson 1979) of a crystalline raw material.

I also agree that reflexivity “resolves the debate” over mediated perception, but I differ with the authors’ contention that “language” [in their terms] derives from depiction and that only such a language facilitates reflexivity.

First, the argument that language is context-independent misses the point. All aspects of culture are dependent on the continuity of a social context, whether in learning to make handaxes or learning a language. As Vygotsky (1986 [1934]:237) pointed out and as recent studies have emphasised, language usage depends upon a series of nested contexts (Bronfenbrenner 1979, Lock and Fisher 1984). Likewise, the meaning of artefacts is contextual; an artefact is, as Mead (1934:280) says, a “collapsed act.” A chair is, at Medicalization of the “frozen gesture” is defrosted. Davidson and Noble’s narrow definitions of “language” and “culture” exclude much of the activity that makes modern humans “modern.”

It is, I think, incorrect to dismiss all comparative study as a “clever-chimpanzee fallacy.” Most primates live socially, and it is at this level that comparison is instructive. Thus, for example, chimpanzees and humans share the ability to recognise themselves in a mirror (Desmond 1979). The recognition of individual identity is central to the fluid social structures observed in both species. Again, primate acts of deception, cooperation, and what de Waal (1982) has called “reconciliation behaviour” indicate that the integrity of social groups is maintained by intentional communication (Byrne and Whiten 1988). These forms of communication are not Mead’s “signals,” because they require what he calls “taking the attitude of the other” [p. 191].

With respect to the role of depiction, it is incorrect to state that non-human primates cannot recognise representations. The Temerlins’ chimpanzee Lucy, for example, derived some satisfaction from Playgirl (Temerlin 1975). Since this occurred only during oestrus, of which humans exhibit no outward signs, it could not simply be an example of imitation.

Primates and humans live in social worlds, and reflexivity, as Mead (1934) and Vygotzky (1986 [1934]) emphasised, is a primarily social construct. It is not founded on a Cartesian relationship between organisms and objects (Humphrey 1976).

The archaeological record gives some indications of the importance of sociality in human evolution. As Davidson and Noble state, the Lower and Middle Paleolithic are characterised by an extreme conservatism in technology (Wynn 1988). Yet most of the radical changes in hominid morphology took place before 100,000 years B.P. (Holloway 1983b), and at the same time most of Europe and Asia were colonised. This evidence suggests to me that it is the construction of society, not the creation of artefacts, which has been the major force in shaping human evolution (contra Foley 1987).

I do not wish to belittle the Upper Paleolithic. My objection is to the ascription of a single locus to the “origin” of “modern” humans and to its identification with the appearance of depiction. It seems to me that consciousness has been continually reconstructed throughout human evolution. The appearance of depictive art is certainly a landmark in this process, but it need be no more significant than other manifestations of social change. The appearance of individually attributed works of depiction in the Renaissance or the birth of McLuhan’s [1964] Gutenberg Galaxy might equally mark the “origin” of “modern” man, if such an origin existed.

JOHN HALVERSON  
Stevenson College, University of California, Santa Cruz, Calif. 95064, U.S.A. 15 x 88

That “depiction transforms communication into language” is a most interesting proposal, but with the best
will and considerable effort, I have not been able to find any argument for it. I do discern a set of parallel propositions: A. Depiction is referential, quasi-permanent (partly by repetition), and context-free and depends for its recognizability on a system of shared meanings. B. Language is referential, endures through repetition, and is context-free, and its intelligibility depends on a system of shared meanings. What I do not see is any causal connection between A and B. [I trust that no one is arguing that because depiction entails \( x, y, \) and \( z \) and language entails, \( x, y, \) and \( z \), depiction entails language.] Even if there were a connection, how could we know which came first? Nothing in the paleontological and archaeological record suggests the priority of either one. Therefore, the argument, if there is to be one, must be logical. It must be shown that logic requires that depiction precede and entail language and, conversely, that language could not arise without prior depictive activity. I see no such demonstration.

Davidson and Noble do lead us along a number of interesting paths, evoking the hope that they will converge somewhere, but the paths often seem very remote and tortuous and have not a few red herrings strewn along the way. The "reflective" aspect of language is often referred to. It is reflective because it is referential; i.e., reference entails reflection. But is this so? It would seem to me that a great deal of normal speech is anything but reflective, even in the minimal sense of simple consciousness of speaking. When such reflection does occur, as of course it may, it is "rendered," as we are told, "by means of language." The verb "render" has many meanings, but I take it from what follows that the essential claim is that reflection entails language: no language, no reflection. But doesn't wordless mimicry involve reflection? Anyone mimicking must surely be aware of what he is doing. The possibility of "wordless awareness" is not shown to be a false assumption. In any case, we have not yet found a place for depiction.

This appears to be sought through the ability to see resemblances. Davidson and Noble assert that it is [only] the making of things to be similar to others (i.e., depiction) that provokes or facilitates the seeing of resemblances. But they also say that the capacity to see resemblances already depends on reflection, which can only be "delivered" through language—"or at least some system of simulation." Leaving aside the final qualification, are they not saying here that language precedes depiction? Leaving in the qualification, are they not admitting that language, "as we know it," is not necessary to seeing resemblance? That depiction might facilitate seeing resemblance is possible, but I see no argument for it, nor do I see how it might transform communication into language.

"Seeing a resemblance" may have at least two meanings. One is seeing that this [real] bison looks like that [real] bison. Another is seeing that certain lines and shapes look like a bison. I presume that the first sense is not at issue here, since the ability to categorize is common and essential to virtually all animal life, and that it is in the second sense that the problem arises, where seeing [taking note of] a resemblance presupposes a perceived distinction between an object and something else that looks like it. But the latter need not be a depiction or any kind of artifact. Davidson and Noble affect a certain scorn for mental imagery, but it is hard to believe that early hominids did not dream or have memory images, just as we do. [Not even Coulter denies the reality of such imagery; in fact he explicitly "in no way contradicts the mundane experiences of . . . mental images or dreams" [Coulter 1983:127; cf. 138.]] Mental images could well have provided the earliest distinction between object and representation. So perhaps could certain natural physical configurations. And so, certainly, could mimicry, especially depictive gesture (drawing shapes in the air), for to be meaningful such [unfrozen] gestures would have to be seen as resemblances, with an attendant distinction made between gesture and referent, and such acts would have preceded depiction—all by the authors' own hypotheses.

In a similar vein [I think], discussing depictive and nondepictive "traces," they assert that for traces to have meaning they must be made, as distinct from just left. Meaning for whom? Left traces obviously are meaningful to a tracker. To say that "once a footprint has meaning within a system of reflective language, any footprint can have that meaning, and so can a stencil . . . or a drawing of a foot" seems to beg the question. Why cannot there be such meaning without language—or depiction?

"Depiction in particular provokes the reflectivity that in turn permits referential utterance." This seems a nutshell summary. Of course I would not disagree with the first clause, but what does the rest of the sentence mean? That reflection is a necessary condition for language? And that depiction is a necessary condition for reflection? This does seem to be the authors' position. But "in particular" does not mean "solely," nor does "provokes" mean "causes." Deixis is a good, traditional candidate for the transition from call and gesture to speech and, though definitely referential, does not imply reflection. If I point and give a warning shout, there is no reason to suppose that this activity is accompanied by any, let alone linguistic, consciousness that I am pointing and shouting. Reflection, if it comes at all, will come after the act, a consequence of, rather than a condition for, referential utterance. A similar fudging occurs in "depiction . . . must be embedded in a system of shared meanings . . . itself expressive through syntactically controlled utterance." Of course it may be so expressed ("expressible"), but that does not mean it is or need be so expressed.

On the related issues of imagery and memory, the authors are led to implausible assertions and odd arguments. For example, from the observation that eidetic imagery (a rare and even rather doubtful phenomenon in any case) is not found in children under six, they come to the transparently specious conclusion that imagery perse, any imagery, "is a property of children's cognition only after they are competent language users and pictorialists." Again, a "keystone" of the argument is that
depiction, by fixing images, "creates memory." Are they really denying memory to predepicting hominids or to living primates? Vygotsky argued for the development of a different kind of memory based on the use of signs—quite a different matter, surely. Another part of the "keystone" is that depiction "allows copying, transporting, and recopying," and this "starts to look more like language." It might begin to look like written language, but it certainly does not sound like speech. Yet "this is the way in which communication was transformed into reflective language." If there is an argument here at all, it eludes me. Again, from Olson's studies of young children, they conclude, "It is only when the writing of text has come into existence historically that there can be a difference between the given and the intended in an utterance." (Apparently lying was unknown in the golden age before the advent of writing.) In fact, Olson's studies do not pinpoint literacy as the decisive factor in the ability to distinguish form and meaning in utterances, as he himself has acknowledged (David R. Olson, personal communication, 1988). But what has any of this to do with the subject anyway? It is the basis of an analogy: just as preliterate children do not distinguish between what is said and meant, "pre-depiction humans did not see the distinction between seeing a thing and responding to it." This is not an argument, nor is the idea credible in itself. Why should we suppose that everything Palaeolithic people saw engendered a response? If it is implied that they only "saw" what they "responded" to, the argument is circular. If seeing itself counts as a response, there is no distinction between pre- and postdepiction people.

It is perhaps the authors' animus toward any hint of "mentalism" that leads them into such dubious ways. Though "mind" may be a modern concept (though not so very modern, since Plato and Aristotle had it), it does not follow that whatever this shorthand expression refers to did not exist before we had the word and began to talk about it. I doubt that anyone in the Palaeolithic had the concept "reflection," though Davidson and Noble freely attribute reflectivity to them. Is this not, in fact, an egregiously "postliterate" concept? On the whole, although the paper certainly provokes reflection, I don't think that its principal thesis, that depiction transformed communication into language, has in any way been plausibly explicated or demonstrated. But I admit that I found the whole paper so mind-(or whatever) boggling that my perceptions of it may be faulty.

GORDON W. HEWES
Department of Anthropology, University of Colorado, Boulder, Colo. 80309-0233, U.S.A. 14 x 88

The advent of durable representations, mainly of hunted animals, in limited regions during the Upper Palaeolithic was only one symptom of significant cultural change, associated with more widespread changes in technology. The shift to anatomically modern Homo sapiens had been under way for a considerably longer time. I agree with Davidson and Noble that these phenomena also had something to do with language. I disagree with their notions that "depiction" was something entirely new and that it played the kind of cognitive transforming role in the Upper Palaeolithic that they attribute to it. Depiction (as they acknowledge) is an essential aspect of gesture, and along with deixis it lies at the root of language. For most of the time that hominids have used language, depiction persisted not in lapidary or other markings but in memory and oral tradition. Some cultures have survived into recent times with far less depiction in material media than that which we have from the European (and Siberian) Upper Palaeolithic. Are members of cultural groups with little or no such material depiction (outside of gesture, dance, mimicry, etc.) cognitively deficient? The notion that "primitive" peoples cannot recognize photographic depictions has not been supported by fieldwork (see Eibl-Eibesfeldt 1972).

Davidson and Noble ignore analogues to depiction which have arisen through natural selection in the "mimicry" of various animal species. That highly intelligent, tool-using hominids would eventually have begun to create cultural equivalents of biological mimicry is not surprising. There is abundant ethnographic evidence of depictive mimicry in connection with hunting and fishing, often involving the use of animal skins, feathers, horns, etc., for disguising hunters or decoying game. Few such devices are made of durable materials such as stone, clay, or bone.

The authors recognize the significance of the decoding of animal tracks by hunters and that we are the only species which does so by responding visually to the shapes of such tracks. I see no perceptual difference between the decoding of animal tracks, not produced deliberately, and the decoding of man-made "depictions," although the effective reading of hoofprints may demand greater cognitive ability.

Davidson and Noble accept the findings of an experiment by Winner and Ettlinger (1979) with two chimpanzees, purporting to show that apes do not respond to photographs. Most of those who have worked with chimpanzees in language experiments would disagree. Premack (1983:chap. 6) reports that chimpanzees do not acquire picture recognition until about age six years, whereas human children do so by two or so. Language-trained chimpanzees perform better than untrained ones. This whole question requires much more research.

The remarkable concentration of Upper Palaeolithic depictive representations in limited geographic areas suggests not a direct connection with language but the existence in limited regions of specific cultic traditions, presumably related to hunting and sexual symbolism. Language, if not as fully developed as in later times, was almost certainly an already ancient part of human cultural behavior and present throughout the human geographic range, in most of which depiction in durable media is absent. If we restrict "depiction" to the kinds which survive from the Upper Palaeolithic, many ethnographic groups are without it to this day, although they possess complex languages and oral traditions replete
with "depiction" in the literary sense. Depiction in this broader mode is a central attribute of language, which characteristically builds "images" of the real or fictitious world. It is also found in dance and ritual and in the use of animal disguises and decoys mentioned above.

It is not necessary that depiction exist only in a matrix of meaning shared with others. Idiosyncratic depictions, whether verbal, gestural, graphic, or plastic, are commonplace, although the materials of which they are composed are usually derived from the surrounding culture. The observed deficiency of apes with respect to depictive behavior probably results from more than a failure to share meanings, linguistically coded or otherwise. The elaborate experimental procedures employed by Premack (1983) suggest how difficult it is to investigate such matters in living apes. Getting into the minds of our Paleolithic ancestors is almost certainly far more difficult, although I do not recommend that we abandon the task.

Reply

IAIN DAVIDSON AND WILLIAM NOBLE
Armidale, N.S.W., Australia. 30 XI 88

We are grateful to the commentators for the care with which they have responded to our paper, and we welcome the opportunity to make plain some issues that previously were not.

Armstrong recognizes that there is continuing speculation that the Upper Paleolithic is a time for the origin of language. Along with others [e.g., Miller 1981, Bradshaw 1988], he does not really address the theoretical issues attending that speculation. In his view we are vague about the nature of the precursor to language. The punctuational argument does of course require some prior development of what he calls "language-like abilities." Natural selection must have some variation to act upon, and we can hardly believe that that was something entirely unlike language. Among other things, we have been trying to identify some aspects of what it is that distinguishes the merely language-like from language.

Armstrong is wrong in suggesting that we require all societies with language also to have depiction, and in non-perishable media. In our argument depiction is essential only in the development of awareness of meaning. Once there was this awareness, meaning could exist without depiction, and any group with language would have the capacity for depiction without necessarily exercising it. Depiction is required in the process of the emergence of language but not afterwards. Further, we argued that, from the evidence of the archaeological record, it is unlikely that the earliest depiction was confined to perishable media. It does not affect that argument to observe some subsequent cultures employing only perishable media.

Armstrong implies that we take the depictions of the Upper Paleolithic of Europe to be the ones that made language possible. In fact, we assert that the earliest claimed depictions [the "vulvas" and Vogelherd figurines] are not the earliest.

By way of critique, Armstrong relies on Premack's (1986) argument that the "linguistic capacity to support the technology of the 20th century existed but was not revealed in the material culture for many thousands of years." From this he takes it that Lower Paleolithic technology [which "in no way required language"] can be explained as having this character without disturbing the view that "linguistic capacity" was present. We prefer a more modest approach to the business of reasoning from the evidence. If there are no signs of language in the Lower Paleolithic, why argue that the beings in question had language but, in effect, "kept quiet about it"? In our view it might be more profitable to consider the Premacks' (1983:29) comment on Sarah's failure to ask questions spontaneously: "The ape's failure is due to its inability to recognize the deficiency in its own knowledge. Language training can supply the form of the question, but it cannot teach a creature to examine the state of its knowledge or to find deficiencies that impel the desire for information." The conservatism of Lower Paleolithic technology seems to us consistent with a failure to question the potential it presented rather than a suppressed ability to do so.

Armstrong also invokes Premack's (1986) notion that language is "more powerful than one can account for in terms of selective fitness." This statement depends on a premise that natural selection creates variation. Selection operates, however, on existing variation—hence there is no predicting the "power" of new variation. There is no denying that "powerful" variations might confer selective fitness, but how can anyone claim to know what is or is not the requisite amount of "power"?

In our scenario, the absence of language accounts for the relative stasis of the Lower Paleolithic, and the selective advantage of language can be seen in the dominance of Homo sapiens sapiens and the generation of unprecedented human diversity after 50,000 years ago. The power of language, given our account of it as essentially reflective, is in permitting a new source of variability upon which selection can operate, and this ultimately includes reflective variability of language itself.

We reject absolutely Black's assertion that the "sequencing of development of structures prior to development of functions... will be found wanting." The emergence of structures before functions follows from the point that selection operates on existing variation. Incidentally, we did not offer a scenario in which "the Cro-Magnon artist" makes a "conceptual transition" from imitative bodily posturing to a reproduction of an animal by drawing it. Rather, we talked of a creature's making a gesture that traces in the air the distinctive shape of an animal, subsequently leaving a persisting trace, on or in a surface, of that gesture. Black is not the only commentator to have had difficulty with this point.

Calvin, while agreeing about the significance of depiction, suggests that depiction abounds in Lower...
Paleolithic hominid groups. The story he tells has much intuitive appeal, but the capacities relied upon—for seeing resemblance, for reading animal tracks, for leaving deliberate signs such as a dragged footprint to indicate an action to be followed—all entail precisely the capacity for reflective representation that we think need to be accounted for. The seductiveness of his scenario lies in its formal differentness to contemporary linguistic practices. We observe, though, that the simple-looking “charades” he describes rely, in contemporary cultures, on already established linguistic capacity—on the capacity to invest a mark with agreed-upon meaning. We would be more convinced by his argument if there were documentation that any non-human animal, primate or non-primate, is capable of reading such signs unambiguously, let alone creating them in an innovative fashion.

What we are faced with in Calvin’s account is a version of the “clever-hominid” theory, in which hominid forms were naturally selected for that were “shrewder” than others. We suspect that this sort of argument is not invoked with respect to the evolution of species other than the human. We see it as deriving from human cultural selection [see Rindos 1986], in which, by means of access to restricted information or other resources for more effective planning, fully lingual competing groups may “outsmart” each other to establish dominance among themselves. A response to this point might be to suggest that Lower Paleolithic hominids were already cultural creatures [see, for example, Holloway 1969], but we find no support in the archaeological record for this suggestion. There is little evidence that Lower Paleolithic artefacts were made and used in the innovative ways that it would require. The evidence adduced, including spurious claims for aesthetic appreciation of fossils [Davidson, Sutton, and Cook n.d.], is better attributed to archaeologists’ need to see design.

Arguments of the sort advanced by Calvin [like de Waal’s 1982] story of chimpanzee “politics”) are projections from current human life. They are instances of what Coulter (1983), borrowing Quine’s (1960) terminology, calls the “transparency/opacitv elision.” A transparent ascription, in this context, is one person’s description to another of a third party’s behavior; it is a description “true” for an observer. An opaque ascription is a description of the behavior from the perspective of the party in question, a description “true” for the actor. Some anthropologists mark this distinction with the terms “etic” and “emic.” A common problem in social science is the elision of these two kinds of ascription, taking it that what is true for the observer is true for the actor. In human communities it is, of course, possible in principle to check whether these two ascriptions coincide—whether the elision is justifiable. In the case of creatures from whom an account cannot be elicited, this procedure cannot be followed, and the elision may occur unchecked. [The terms “transparency” and “opacity” are doing quite different duty in this context to that in our article, where they are used with respect to iconicity vs. arbitrariness of signs.]

While Davis agrees that the “cognitive” effect of depicting is profound, he believes we are wrong in saying that it requires a shared system of meanings. But meaning need not be universally or timelessly shared. An item’s meaning may be limited to a tiny and transient subculture (witness family nicknames). It must, however, be the property of more than one member, and it must be repeated; this is nothing more than a logical property of any language or system of representation. The image that Davis presents of the individual alone with her or his meaningful depiction is one allowed only within a culture that sanctions individuality of expression. It is logically impossible for an individual to hold a meaning privately [one does not by this mean undisclosed], as Wittgenstein showed. For meaning to be meaning it must be shared. Contemporary philosophers can imagine private meaning because their language games permit them to persuade each other of [to give meaning to] its imaginability. It can correspond to no pragmatic arrangement of persons.

As to ambiguous meaning and the way meaning emerges within cultures, we do not believe that we are at odds with Davis in holding that meaning is identifiable from the ways in which objects are used, hence the ways in which their referents can be intelligibly incorporated into utterances. Ambiguity no doubt arises from many sources, including plausibly the incorporation of a use of an object at the margin of the culture by persons not fully tutored in its sanctioned use. Modern people exposed to “Western” post-Renaissance culture can understand the ambiguity in a depiction [a cube that looks like a pyramid] because of traditions of perspective drawing that support both specifications. The capacity to see that ambiguity cannot be construed as solely an individual act of creative meaning-making. Girgus, Rock, and Egatz (1977) have shown that ambiguity may even need to be a topic of instruction before it can be perceived in a picture.

Our argument is that the capacity for making marks or traces arises directly, perhaps at first accidentally or incidentally, from the capacity for manual gesture. The “macaroni” at Pech Merle (Lorblanchet 1988) or Altamira could plausibly have been made by hand movements without any gesture of mimicry [rather as similar marks were made by apes trained by Morris [1962]]. If we postulate a creature that cannot see resemblance and has no gestures that mimic features of its environment, there is no mechanism by which its observing these squiggles can lead to their having meaning or resemblance. Squiggles made with no context of imitative gestures are traces left but not made. An accidental or incidental squiggle could result in a trace that is seen to resemble a bison only by virtue of the existence both of bison and of aerial gestures that follow their characteristic outline. Borrowing from a remark made to us by Peter Jarman, we could go further to say that the accomplishment of “perceiving resemblance,” in such acci/incidental circumstances, could derive from whatever performed trace happens to conform with the contour of the gesture, as if stroking the hump of the beast. This action would mesh a novel appearance [the
trace] with a familiar visuo-motor event [the gesture, or
strok[ng the hump of a dead bison], allowing their
morphological kinship, their shared characteristics, their
"resemblance" to each other [and to the object of the
gesture] to be perceived.

Davis asks why pointing would not be as powerful as
imitative gesture. Pointing seems to be uniquely hu-
man: the Premacks [1983:56] concluded that although
chimpanzees could be conditioned to respond to and use
it they did not do so spontaneously outside their test
space. We see pointing as standing between the one-
handed aimed throwing to which Calvin has called at-
tention and imitative gesture. For a conspecific observer,
the orientation of hand and arm toward the object in
aimed throwing becomes salient just because of what
happens in connection with that orientation. In aiming,
the arm from which the missile is launched must point
directly at the object at some stage in the throw cycle.
As a first suggestion, aimed throwing entails prior align-
ment of throwing arm with target, hence the successful lia-ison between aim, fire, and strike provides the pre-
paratory aim with the function of orienting an appropri-
ate observer (one also able to perform this act) from the
arm to the thing aimed at. But the argument works even
without prior alignment. At the moment of launch in
natural aimed throwing (by which we mean throwing
uninfluenced by the conventions and rules of athletic
games) the arm/hand must be angled directly at—
pointed at—the target. Thus pointing, as an act that
orients an observer from the pointing arm to the location
pointed at, may originate in one-handed aimed throw-
ing.

Now, gestural imitation entails being able to extend
pointing into what can be construed as an outcome of
prolonged pointing. The capacity to pursue such a path-
way of pointing is delivered by the sequencer control
system. Imitative gesturing [outline pointing] requires the
very kind of visual-motor control of the limbs and limb parts in relation to each other and to the torso that
is part and parcel of aimed throwing. As to why a cre-
ture with the capacity to control its limbs in this way
might perform a gesture in outline of its target or some
other salient creature, we confine our speculation to the
notion that acts are performed because they can be
rather than seeking any adaptive or utilitarian function
presupposing intentionality. The employment of the
body in exploratory engagement with the world—what
can be called "play" in human cultures—need carry no
cargo of intentionality. What does need impressing into
our account is the idea of salience. A gesture might, in
theory, be considered capable of indicating any number
of things. But in the social circumstances of spotting a
bison—a large, mobile, and self-contained entity—the
gesture that has the effect of drawing attention to its
presence is a response to an event of high salience. As
does the worm for the toad, the bison picks itself out of
the background by virtue of its size, mobility, and unity.
The gesture arises by virtue of capacities previously se-
lected for.

An argument for the salience of creatures like the bi-
son is supported strongly by the archaeological record—
which, although later than the ancient events we are
considering, is still very early—just because it is a record
of depiction of animals and fellow creatures, of the ani-
mate world. The landscape is not pictured. We will in
due course be exploring the idea that the prehistoric
post-depictive "seeing" [vs. pre-depictive "responding"]
that we have been arguing for is initially limited to the
"seeing" of the animate environment—that the "see-
ing" of stationary features of the landscape is a subse-
113] appear to have had more success in getting
chimpanzees to respond to videotapes than to single still
photographs. Gibson [1979] argues that vision evolved
for the detection of movement, whether of oneself or of
independent environmental features, and Cutting [1986]
has recently formalized that position.

The issue that remains to be considered is the selec-
tive advantage of imitative gestural signalling. Just as
chimpanzees must learn to be silent [see Kortlandt
1973), we may recognize the selective advantage of
aerial outlining and imitative posturing in silence for a
relatively defenseless creature indicating salient features
of its environment to conspecifics. We need not, in other
words, read intention into the picture just yet.

What we are calling the "sign" of the gesture—its
record, as well as its reference—is what introduces the
denotative ("this thing can stand for that thing") ele-
ment. One cannot point to a gesture, but one can point
to its result, the trace, and one can point to that just as
one can point to the real thing. And one can return and
point to the depiction and see the relation between it and
its object by virtue of their visual similarity to the
iconic gesture. We think that this is the moment of
achievement of reference and is what constitutes the
perception of resemblance. The noticing of resemblance
between x and y, between a cloud and a clown, is a post-
reflective achievement, arising from the capacity to rep-
resent to oneself and others—that is, to comment
upon—the nature of one's perceptions. Davis, having
seen the importance of the origin of the seeing of resem-
bance [1986a:199], waves away the real problem in say-
ing that such seeing can be taken as a "primitive."

Davis [1986a] made no assumptions about the pres-
ence or absence of language in the early history of image
making, and he continues to view the origin of depiction
as independent of the emergence of language. Our con-
cern is with the emergence of language; the interest we
have in common with Davis is in the prehistoric origins
of things symbolic, hence distinctively human. We
think that their origins are inevitably related; he does
not.

We are still hesitant about the place of intention in our
account. It makes perfect sense to us that behaviors arise
because the creatures concerned have certain evolved
capacities. With the emergence of the capacity to depict,
a practical power has evolved—a means for providing
reflectivity of product on producer, the making of a rec-
ord, hence a memory, and yet an image that resembles
its object—that provides a (the?) direct means for re-
ferring, for indicating an object in its absence. No deliberate, i.e., self-conscious, intention need be assumed in all this. With further exploratory actions in relation to the new power, it will be perceived that the very act of tracing has malleable or variable properties. The trace can be made to resemble other things, or to resemble nothing. Eventually, the concept of depicting will arise as those making or seeing the images comment on the depictions themselves as depictions. All this will be exposed to selective pressures, because this new behavior introduces more of the variation on which natural selection operates. Moreover, this form of variation allows reflection upon itself and upon the selective pressures themselves. If depiction enhances the ability of some creatures to communicate about the objects they have depicted, then there will be selection for depictors over non-depictors and for “good” depictors over “bad” ones, provided that communication is important for survival. We treat very lightly at this point because we do not know what it is like not to be able to name things, not to know that, not to be a conceptualizer. While we think we can account for the way in which the making of depictive traces could have originated, we are not seeking to explain the origin of the concept of depiction. A concept is a cultural invention and therefore follows from the emergence of language. We cannot tell from a thing we take to be a depiction whether such a concept existed.

Falk simply restates the “agreed-upon paleoneological” position that we are necessarily at odds with. It may be the case that language emerged slowly and steadily, as suggested by the hominid skull evidence. Where is the concomitant evidence of increasingly lingual forms of life? Falk asks what the stimulus for brain-size increase might have been. Leaving aside the issue of alomorphic increase in brain size as body size increased [see, for example, Holloway 1983], as there is certainly a circularity in any argument about causality there, we agree with Calvin (1982) that it has to do with selection for the use of skills such as aimed throwing and manipulation. We think it likely that there was also selection for more complex (“language-like”) communication, which is why we think that it would have been easier to teach language to a Neandertal than it is to teach it to a chimpanzee. But the issue we address is how that “language-like” communication was transformed into language. Falk (1980a) herself once suggested that the issue was what the australopithecines had to talk about. We are saying that it was the evolution of “talking about” which characterized the origin of language. We might argue against Falk and the other paleoneurologists that if early hominids already had language when they had small brains, it seems difficult to imagine what the selective pressures for larger and more complex (see Holloway 1983) brains were. After all, one of the distinctive characteristics of linguistic capacity is the increased potential for non-genetic variation on which natural selection can operate.

Foster incorrectly imputes to us the fallacy of assuming that language equals naming. Even if we had made such an assumption, however, it fits fairly well with Foster’s report that the modern child’s introduction to language is largely on the basis of nouns. Language functions in the accomplishment of many practical ends. What we are pointing out is that the capacity to refer, be it to objects, events, or their conjunction, is at the heart of its reflective character. Naming is an essential element of language, though not its only element.

The bits of Upper Paleolithic art which Foster cites are all later than our suggested origin of language and therefore capable of having been produced with symbolic code. The claim that there is a directional change in Upper Paleolithic art from opacity to realism is specious, and, aside from the “vulvas” as earliest specimens, examples could be selected which showed the opposite trend [Vogelherd horse to Parpalló geometrics or Gönnersdorf female figures, for example]. The “opacity-to-realism” trend has received most emphasis from scholars determined to write a chronology from the depictions by fitting them to some sequence about the evolution of representation. Part of the problem arises from Foster’s observation about the incompleteness of early representations. Implied here is that a representation can be complete, whereas it is the essence of a representation that it cannot.

Both in suggesting that “the earliest known art begins with . . . opacity and progresses toward greater realism” and in her account of early child language, Foster is expressing faith in indissoluble links between the evolution and development of language and the evolution and development of depiction. Our disagreement on this point is fundamental, since, in our account, once language has appeared the evolution of the two systems can be autonomous. Part of her argument, however, confirms our position. She asserts that vocal utterance originally replicated “movements and spatial relations in nature,” and while she chooses to characterize this as “abstract” it is evident from her own description that we are looking at a change from iconicity to arbitrariness.

Graves says that we have neglected the social basis of linguistic meaning. We simply take it as given, however, that, as Mead emphasized, meaning arises in social, not individual, forms. Our project is to find a plausible means whereby utterance was invested with referential meaning. When monkeys make warning cries in the presence of predators and conspecifics, it has meaning for the colony insofar as it signals “danger-from-a-certain-direction.” But these creatures accomplish no reference outside of this context such that the form of reference itself can be perceived as carrying the meaning. [This, incidentally, is what we mean by context-independence. Of course we appreciate that meanings hang upon the continuation of social contexts; why else would we insist that meaning is socially shared?]

Higher primates are by Graves’s own account “social” creatures. Why, then, do they not have language? Hominids were also “social” in this sense, but it almost looks as if their colonization of Eurasia was made possible by fire and handaxes while natural selection shaped their morphology—an account that attributes no special im-
portance to their social interaction. Moreover, not all forms of interaction that ethnologists would call social are identical to human social behavior (e.g., Hinde 1974: chap. 2), and only at the general ethnological level can we say that they were important amongst hominids in earlier stages of evolution. Social interaction in a human sense has surely played a vital role in human evolution since the time we became human. To conflate these categories of interaction is to assume one's way out of the problem of how one form evolved into the other.

Graves opposes a single locus for the origin of modern humans. How can an evolutionary argument cope with multiple origins of an organism as complex as "modern" humans? The argument we are making accepts that there was a polytypic species, H. erectus, in Africa and Eurasia by at least 0.75 million years ago. Natural selection operated on one of its populations, probably in Africa, to produce changes in the skeleton which made those hominids look more like "modern" humans than any other contemporary hominid. We cannot know what advantage was conferred by that shape and what it can do, but it can surely be argued that that advantage allowed "modern" humans to replace H. erectus elsewhere. Although this version telescopes or bypasses many debates in physical anthropology, an evolutionary argument must take this form. Graves writes of consciousness as having been "continually reconstructed throughout human evolution." It may have been so throughout history (Mead 1932), once it had emerged through reflectivity of reference; thus writing, for example, may have been independently constructed within different cultures. Surely, however, this cannot be so of something as fundamental and ultimately biological as consciousness itself. It may be that Graves is uncomfortable with the necessity to incorporate language and "consciousness" into evolutionary arguments.

Halverson's position is philosophically opposed to ours. He does recognize, in common with us, that the force of the argument must be the theory or logic. He argues that "anyone" who mimics must surely be "wordlessly aware" of what he is doing. We agree that any modern adult human being who mimics will be aware (though not wordlessly) of her or his actions, but our discussion is about hominids. Given that he finds it "hard to believe that early hominids did not dream or have memory images, just as we do," it seems likely that he considers them already human and therefore cannot address the problem of how behavior which was distinctively human evolved from something which was not.

Halverson sees a circularity in our speculation that making something that resembles something else provokes reflectivity but "seeing resemblance" (which he presumes must precede the making of the resembling thing) depends upon reflectivity. The elaboration we have offered in response to Davis's related criticism seems applicable here.

Halverson follows Davis in proposing two senses of the term "seeing a resemblance": one in which the organism sees "that this [real] bison looks like that [real] bison" (which he takes to be common to virtually all animal life) and one in which "seeing [taking note of] a resemblance presupposes a perceived distinction between an object and something else that looks like it." He identifies the second as the basis of representation, wherein mental images, natural features of the environment, and depictive gestures could all be seen as resembling other things, and be so seen prior to any depiction. But these two senses of "seeing resemblance" are not really separate. To see that x resembles y, whether x and y are real objects or x is an object and y is "an image" of it—is, as Wittgenstein (1958) pointed out, an achievement, not a process. To "see resemblance" is to make a claim about one's perceptual experience and depends upon the ability to comment on that experience. That an animal can be observed to respond in the same way to similar objects does not amount to evidence that it "sees a resemblance" between them—that it sees that "this [real] bison looks like that [real] bison." To see in this way entails "taking note of" one's experience, and we do not know how that can be done non-linguistically.

We do not "scorn" mental images. Rather, we argue that states of "mind" taken for granted in modern human life cannot be simply assumed in ancestral forms as Halverson would wish. Armed with such assumptions, of course, the whole issue of how humans evolved in the way they have becomes at once straightforward and utterly unfathomable. Mind, memory, imagery, the ability to point to things are all simply and somehow "there" in Halverson's picture of hominid existence. Nothing needs to be explained, and nothing, therefore, is explicable.

Hewes seems to have made the same mistake as Foster of thinking that we mean that the first depictions were the European Upper Paleolithic ones. As does Armstrong, Hewes takes it that a link between depiction and language must persist after the emergence of language. We would not disagree that most pre-literate language has been in oral traditions and not picture stories. Once again, the only relationship between language and depiction in our account is in the process of emergence of language. We agree with Hewes that the Upper Paleolithic pictures were later than language's emergence; we are more circumspect about the motivation for their production.

Hewes misses our point about tracks. What is implicit in our account is that the ability to recognize tracks may appear relatively late in human evolution. Disagreement with this will be common among those who think that there was non-opportunistic killing of animals before the Upper Pleistocene. Similarly, Calvin has hominids hunting far earlier than good evidence suggests. In our argument tracking follows other reflective pattern recognition rather than creating it.

In summary, the major issue for the archaeologists among the commentators is the identification of the signs of language in the material remains of the past. Palenigneologists such as Falk rely primarily upon the conformation of the outer surfaces of the brain as recorded on endocasts of fossils. The more ambitious among them link this evidence with the evidence for
ston tools [Holloway 1969]. We have offered an alternative explanation of the form of Lower Paleolithic stone tools. We do not believe that this evidence speaks unequivocally of language. Hewes and Calvin differ with us about the evidence for such prehistoric activities as hunting and tracking. We do not accept that there is unequivocal evidence for these activities earlier than a period in which there were also depictions. We see no sign in these comments that anyone else can account for the sporadic occurrence of non-perishable items marked or modified non-iconically followed by the appearance of depictions such as the “vulvas” and the Vogelherd figurines. We offer the prediction that there is a body of undiscovered depictions earlier than those from the Upper Paleolithic of Europe.

Differences about the theory of perception and language derive from different philosophical positions about mind [Halverson] and the possibility of its evolution [Graves] or about language and structuralism [Foster]. Our enterprise has been to explain human mind in an evolutionary framework, and therefore we have relied on a “realist” theory of perception and action and a pragmatic approach to language. While we have found Calvin’s argument about aimed throwing most stimulating, especially as a counter to the standard argument of other paleoneurologists, his speculations about hominids’ capacities for observing their environment lack the very kind of theory that we are trying to develop. Amongst our commentators Davis continues to be close to our theoretical position, although he is more concerned with depiction than with language.

References Cited


VON VALLADAS, VUCKO, VUCKO, TRINKAUS, THIEL, TENNANT, YOSEF, bees.

RENT bridge 292:43-56.

Edited the of Anthropology 28:356-41. [DF]


