

## The Distributed Cognition Perspective on Human Interaction

*Edwin Hutchins*

What comes to mind when a social or cognitive scientist thinks about “human interaction”? The answer surely depends on the scientist’s field of study and most of us learn images of interaction implicitly as part of being socialized into a scientific community. In some corners of artificial intelligence, the prototypical interaction is a sequence of turns in which strings of characters or symbols are exchanged. For some conversational analysts, the interactions of interest are mostly verbal, telephone conversations, for example. Ethnographers of speaking may focus on face-to-face interactions, and that formulation draws our attention to facial expression in addition to verbal behavior. To go further in this direction, the phrases that describe our default images of interaction become awkward. Many of us speak about “multimodal interaction,” but at the workshop leading to this volume, Emmanuel Schegloff reminded us that this phrase is ~~an oxymoron~~. So, shall we simply say “interaction” and hope that others’ imaginations are as rich as our own? My personal preference is to emphasize the way participants to an interaction coinhabit a shared environment. No matter how they are described, our default images of human interaction have powerful consequences for the way we do science. Such images guide decisions about where we look for evidence concerning the nature of human interaction. They shape our understandings of what the observed evidence means. (What is the nature of human interaction, and what phenomena are our theories supposed to explain?) Finally, such images affect how we chose to explain the origins of contemporary human interaction.


This chapter has three substantive parts. The first part describes how the distributed cognition perspective directs our attention to particular classes of interactions. The second part uses the examination of an example of real-world human interaction to construct a description of the nature of interaction. This examination shows real-world interaction to be deeply multimodal and composed of a complex network of relationships among resources. It also shows that some cognitive processes are properties of the system of interaction, distinct from the cognitive properties of the individuals who participate in the system. The last part explores the implications of this “naturalized” notion of human interaction for our understanding of both the nature of contemporary cognition and for the kinds of processes that might have given rise to contemporary cognition. What evolves is not the brain alone, but the system of brains, bodies, and shared environments for action in interaction. Cultural practices are as much a part of the story of cognitive evolution as are changes in brain structure. This means that important milestones in cognitive evolution could, in principle, have been achieved without any particular genetic adaptation being associated with them.

### **Distribution means Interaction**

The subfield of cognitive science called “distributed cognition” does not study any particular kind of cognition; it is an approach to the study of all cognition. It assumes that cognitive processes are always distributed in some way. Rather than assuming a boundary for the unit of analysis a priori, distributed cognition follows Bateson’s (1972) advice and attempts to put boundaries on its unit of analysis in ways that do not leave important things unexplained or unexplainable. This means that a group of people working together is a distributed cognition system. In such a case, cognition is distributed across brains, bodies, and a culturally constituted world. Describing the cognitive properties of this unit of analysis has been the most obvious contribution of distributed cognition to cognitive science, and is certainly the most relevant aspect of the approach for anthropologists. An individual working alone with material tools is also a distributed cognitive system, as is an individual working alone without material tools. So too is an individual brain situated in the body, or the brain without consideration of the body because cognition is distributed across areas of the brain. Even single areas of the brain are studied now as systems in which cognitive function is distributed across layers of neurons. And the same is true down to the

level of a network of neurons in the brain. The point is that distributed cognition is not a kind of cognition at all, it is a perspective on cognition. Its chief value is that it poses questions in new ways and leads to new insights.

When applied to systems that are larger than an individual actor, distributed cognition is an approach to cognition that is deliberately framed in a way that keeps culture in mind. When units of analysis that are larger than an individual are examined as cognitive systems, acknowledging the involvement of culture with cognition is unavoidable. Distributed cognition sees real-world cognition as a process that involves the interaction of the consequences of past experience (for individual, group, and material world) with the affordances of the present. In this sense, culture is built into the distributed cognition perspective as at least a context for cognition.

~~Three ways cognitive processes are distributed.~~ From a cultural point of view, cognition is distributed through time,  between person and a culturally constructed environment, and among persons in socially organized settings.

### *Interaction with Social Others*

Just as physical labor can be distributed among persons; cognitive labor can also be distributed among persons. This distribution of cognitive labor is always mediated by human interaction. It relies on human sociality and forms the context for sociality and its development. This was known by anthropologists long before the distributed cognition label was coined. Durkheim and his students, especially Halbwachs (1925), explored socially distributed memory. Douglas's classic *How Institutions Think* (1986) examined the ways that reasoning and rationality are shaped by institutional organization and goals.

The distribution of cognitive labor can give rise to supraindividual cognitive effects. That is, social groups can have cognitive properties that are distinct from the cognitive properties of the individuals who compose the group. Although similar arguments can be made for all cognitive processes, it is perhaps easiest to illustrate these processes in the domain of memory. Jack Roberts's (1964) analysis of the memory storage and retrieval properties of various Native American groups was one of the first to read social organization as computational architecture (although he did not use that language). Bartlett's (1955) seminal studies and theorizing about the reconstructive nature of memory in 1932 led to more recent studies of collective remembering (Middleton and

Edwards 1990). Decision making and interpretation formation are socially distributed in a number of institutional settings including juries, intelligence agencies, military units, markets, and elections (Surowiecki 2004). Computer simulation studies (Hutchins 1995b; Henty 1999) have shown that simply changing patterns of communication among decision makers in distributed systems can change the likelihood of various classes of decision outcomes. The notion that patterns of information flow have cognitive consequences was explored in the domain of commercial airline flight decks by Hutchins (2000) and extended to implications for design of work systems in general in Hollan et al. (2001).

### *Interaction with the Material Environment*

A second kind of interaction that involves the distribution of cognition is the interaction of persons with their material environment. A person in interaction with cognitive artifacts can have cognitive properties different from those of a person alone. Noticing the similarity between artifacts that amplify our physical strength, and those that amplify sensory processes, Bruner et al. (1966) proposed that some artifacts, such as language and symbolic systems, could be conceived of as amplifiers of cognitive capacities. In cognitive science the notion that the immediate environment can be considered as external memory was noted by Newell and Simon (1972). The notion of cognitive artifacts as a class of objects was explored by Norman (1994) and Hutchins (1999).

Cognitive artifacts have their effects by reorganizing cognitive capacities into functional constellations that provide the new capabilities. Cole and Griffin (1980) refined the cognitive amplifier view by noting that these artifacts do not actually amplify any existing cognitive capacity. Rather, when a person performs a cognitive task (e.g., remembering) in coordination with cognitive artifacts (e.g., using paper and pencil) a different set of internal and external resources is assembled into a dynamical functional system (Luria 1966) that does the job. In this “functional-systems” view, cognitive artifacts are transformers of cognitive systems rather than cognitive amplifiers. The focus here is on the interaction between internal processes and structures and processes in the environment. The functional-systems framing of distributed cognition has been applied to flight deck cognition (Hutchins and Klausen 1996) and to the question of how a flight deck remembers speeds (Hutchins 1995b). The latter article showed why a complete knowledge of the psychology of individual memory would be inadequate

to understand memory in an airline flight deck. Humans inhabit a cognitive ecology that contains many sorts of cognitive resources. Some of these are physical objects, some are cultural practices, and some are mental models. Cognitive effects emerge from the interaction of persons with the rich cultural content of the cognitive ecology.

### *Interaction of the Present with the Past*

The development of cultural cognitive ecology is itself a cognitive process. It is a kind of learning process. Culture is a process that, among other things, accumulates partial solutions to frequently encountered problems. Artifacts and practices have historically contingent cultural developmental trajectories. As cultural creatures, we need not discover the solutions to most of the problems we face. Both the framing of problems and their solutions are already available for learning as part of our cultural heritage. Hutchins and Hazlehurst (1991) provide a computational demonstration of the fact that a community can learn things that no individual could ever learn alone.

Evolutionary processes operating in the cognitive ecology can build the structure of a task into the structure of artifacts (Hutchins 1995a). Stated more accurately, cultural evolutionary processes build the structure of task performance into the organization of the system of activity that exists when the artifact is engaged through cultural practices.

Finally, all of these sorts of interaction and distribution take place simultaneously in real world activity. Many of the effects described in this section emerge from the interactions of elements in a complex system. There is an important methodological corollary to this observation. It has been known in cognitive science for some time that behavior patterns that can be economically described by rules or by goal seeking, need not be the products of processes that include the explicit representation of either rules or goals. There are deep philosophical issues here concerning the ascription of processes that include the representations of goals to account for what appears to be goal-directed behavior. (See Clark 2001, chs. 3–4.) The point is to keep the description of the patterned behavior clearly separate from the description of the process that produces the behavior. Many kinds of processes can produce behavior that appears goal directed, but not all of them involve any representations at all. As a simple example, consider Braitenberg's (1984) vehicle number two. This is a simple robot consisting of a body, two laterally mounted light sensors, and two laterally mounted drive wheels. Each wheel is driven at a speed that is a monotonically increasing function of the activity

of the attached light sensor. If the light sensors are connected to the motors ipsilaterally, the robot turns away from light. If the light sensors are connected to contralateral motors, then the robot turns toward light. When observing the robot move in the vicinity of light sources it is very tempting to say that it avoids light or that it seeks light. But words such as “avoid” and “seek” invite the attribution of internal mechanisms, representations of goals for example, that are clearly not present in the robot. The same goes for many other kinds of cognitive processes. That is, many so-called “cognitive processes” are identified by patterns of observable behavior, whereas the nature of the processes that produce those observable behaviors may be very different from the patterns that are produced. This is true at the level of an individual person as well as at the level of groups of persons. What this means for our images of interaction is that once we commit to the notion of rich interaction, even deciding what task is being accomplished may depend on knowing something about how it is accomplished.

### *Real-world Interaction*

All of the sorts of interaction and distribution described in the previous section take place simultaneously in real world activity. This section presents an instance of rich, culturally grounded, real-world interaction. In *Cognition in the Wild* (Hutchins 1995a), I used an extended study of ship navigation to show how the cognitive science of real-world activity could be accomplished. That book emphasized the distribution of cognitive processes between persons and technology, among people, and across time in the development of the social and material context for thinking. My research group recently undertook a reanalysis of some of the video data from the ship navigation study.

### **A Brief Case Study**

I have selected for analysis about ten seconds of interaction in which two navigators (a bearing recorder and a plotter on the bridge of a navy ship) choose landmarks to use in the next position fix (see Figs 14.1 and 14.2). Position is determined by measuring the bearing of landmarks and plotting these bearings on a chart. A plotted bearing defines a line of position (LOP). Three lines of position define a position fix (Fig. 14.3). This is a clear case of distributed cognition. The individual and institutional knowledge of ship's position is produced by the activity of a complex system involving interaction among persons and complex cultural, organized material media.



**Figure 14.1.** Navigation team on the bridge of a navy ship. The bearing recorder is in the foreground. The plotter is to his left.

The navigators have projected the estimated position of the ship at the time of the next position fix (the half-circle in Fig. 14.3). They must now choose three landmarks, such that the LOPs that will be observed at the next fix time will intersect at useful angles (Fig. 14.4). This is an instance of a cognitive process: choice. It is useful to note here that appropriateness of a chosen LOP is not a property of the LOP itself, it is a property of the relations of the LOP to the other chosen LOPs. That is, although a position fix consists of three elements (LOPs), none of the individual elements can be said to be good or bad with respect to the choice criteria. The criteria refer to the relations among elements, not to the elements themselves. This can be taken as a model of a more profound phenomenon to be encountered below. The meanings of elements of multimodal interactions are not properties of the elements themselves, but are emergent properties of the system of relations among the elements.

A transcript of the verbal part of the interaction among the navigators looks like this:



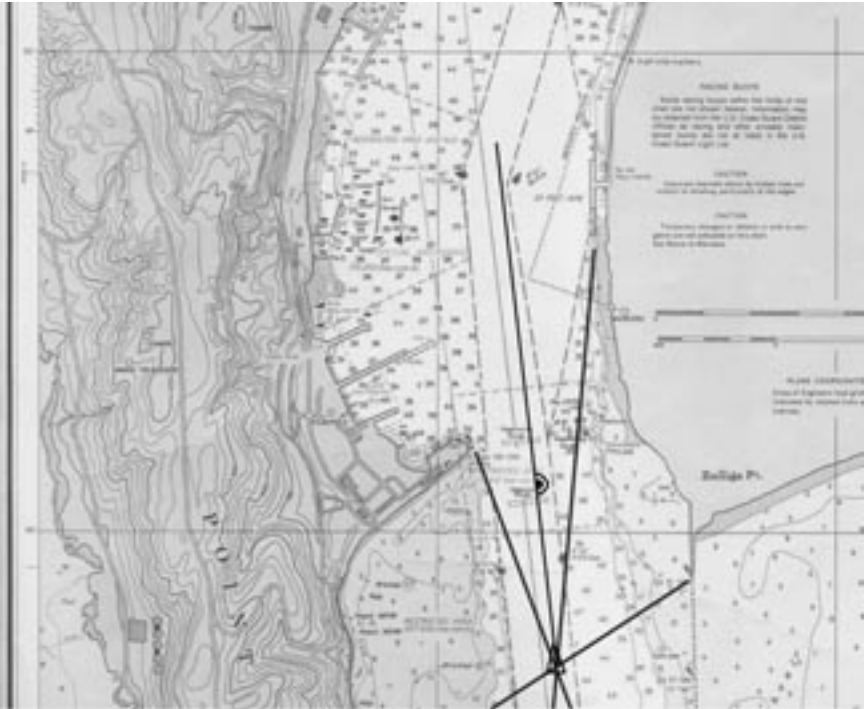
**Figure 14.2.** Enacting provisional lines of position. The bearing recorder is completing his conversation turn while plotter positions his hand to take the (gesture and talk) floor.

- 1 BR: so: it'll be that (1.9) n that (1)
- 2 P: Ballast Point (.7) Bravo (1)
- 3 BR: u:[h
- 4 P: [that's good (.5)
- 5 BR: okay (1.2)

However, the verbal exchange is just one element of the interaction. The next paragraph gives a richer ethnographically informed description of the activity.

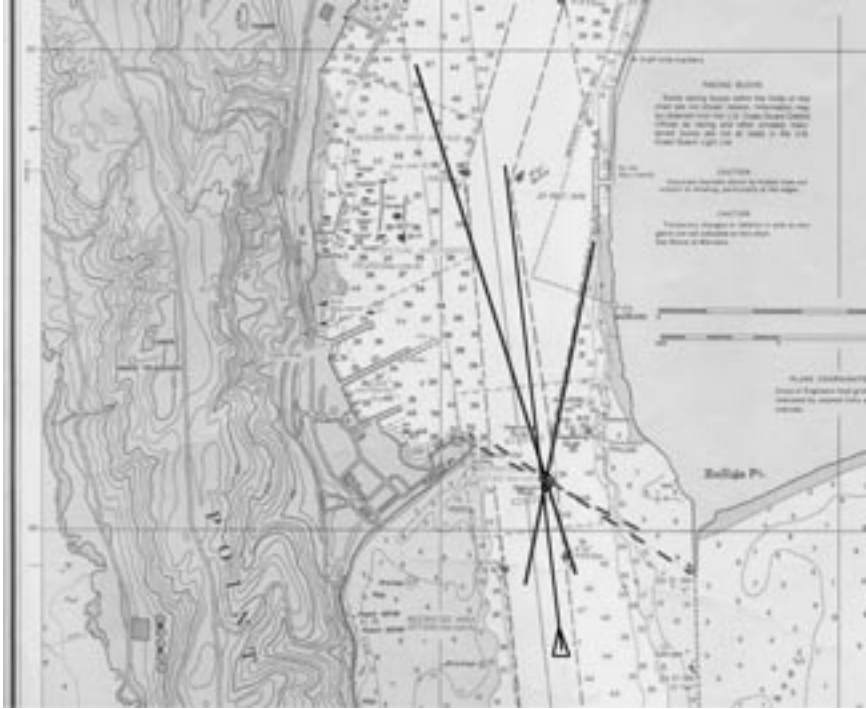
The bearing recorder first proposes two landmarks to use at the next fix. He leans over the chart (saying "It'll be...") and uses his left index





**Figure 14.3.** Three lines of position fix the position of the ship (represented by the triangle). The anticipated course extends from the fix triangle to the estimated position, EP (half-circle), where the ship is expected to be at the time of the next fix. Ballast Point is at left center, Bravo Wharf is above to the right, Light Victor is to the right.

finger to quickly trace a line from a landmark called Ballast Point to the approximate location of the estimated future position of the ship (saying “that”). His finger wavers for a moment making a loose clockwise loop over the chart then he traces a line from the landmark called Bravo Wharf (saying “ ‘n that”). The bearing recorder’s left hand remains in the vicinity of the estimated position and he pauses for one second. (This moment is shown in Fig. 14.2.) The plotter interrupts the bearing recorder’s activity by moving his right hand, middle finger extended, into the area over the chart where the estimated position has been plotted. The bearing recorder withdraws his left hand from the area as the plotter’s right hand comes in. Quickly tracing the imagined lines



**Figure 14.4.** The dashed lines indicate a poorly chosen pair of landmarks for the next fix. The angles of intersection among the LOPs should be open. Using either of the dashed lines with the two piers ahead would produce more favorable angular relations among the LOPs.

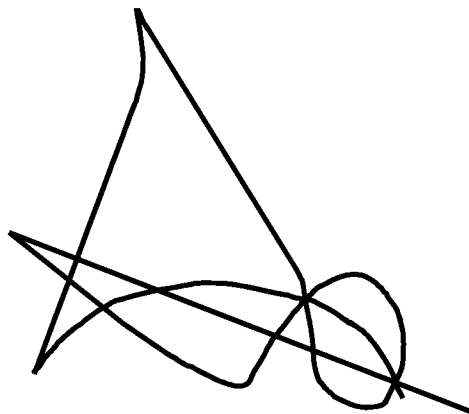
of position from each landmark as each is named, the plotter revisits the same landmarks just mentioned by the bearing recorder, “Ballast Point, Bravo.” The bearing recorder tries to retake the floor by leaning over the chart and reaching toward the plotting area with his left hand, saying, “u:h,” but the plotter rebuffs him by making another gestured LOP from the vicinity of the depiction of Light Victor to the EP (half-circle) and saying, “that’s good.” Because Light Victor is located to the east of the EP, this gesture both indicates a third LOP and effectively blocks the entry of the bearing recorder’s hand to the plotting area. The bearing recorder pulls his left hand back, rests it on the chart table in front of him and says, “Okay.”

As the navigators work, they use their fingers to trace lines from various landmarks to the vicinity of the estimated position. The gestures

enact imaginary or provisional LOPs. These ephemeral structures are the representations on which the choice process operates. The criteria for evaluation are the angles of intersection among the prospective LOPs. The creation and evaluation of the proposed LOPs is carried out in a conversation between the two workers. The conversational turns are multimodal in that they include environmentally coupled gesture, cogesture speech, body orientation, facial expression, and tool manipulation.

### *Environmentally Coupled Gesture*

The gesture is complex. The hands of the participants move around a lot over the chart (Fig. 14.5). Some parts of the gesture stream are meaningful. Some are not. Some gestural strokes represent lines of position, whereas other strokes reposition the hand to begin a meaningful stroke. How do the participants distinguish the meaningful parts of the gesture from the parts they should disregard? First, the participants know that the objects of interest are virtual lines of position. These lines should link landmarks with the ship's estimated position. This is part of the common ground shared by the navigators. As Enfield (2005, this volume) and Clark (this volume) demonstrate, features of a shared task world can contribute to the establishment and maintenance of common ground. The bearing recorder says, "It'll be that 'n that." The seemingly unbound anaphora of "It" refers to the object of the



**Figure 14.5.** The trajectory of the bearing recorder's gesture is complex.



**Figure 14.6.** The trajectory of the bearing recorder's gesture as it was performed over the chart. Tick marks on the gesture trajectory indicate the location of the bearing recorder's index finger in each frame of the video. The filled arrows indicate the LOPs that are made salient by the combination of the many cues produced by the bearing recorder. These are the LOPs he proposes for consideration.

understood current project that is the triplet of landmarks to be used in plotting the next position. An experienced navigator can see the chart as landmarks and EP. The trajectory of the gesture superimposed on the interpreted chart picks up some possible lines of position but seems to have nothing to do with others (Fig. 14.6). The trajectory of the gesture does not unambiguously pick out the potential lines of position that are being proposed by the bearing recorder.

The bearing recorder's gesture also has a velocity profile. That is, some parts of the motion of the bearing recorder's hand are fast, others are slow, and others come nearly to a stop. Velocity is probably an indication of many conceptual elements and of the affective states of actors as

well. Meaningful gestures often come in the form of strokes that are demarcated by pauses before and after the meaningful stroke. These are called pre- and poststroke holds (McNeill 1992). A frame-by-frame analysis makes it possible to indicate the location of the hand in each frame of the video clip. Fig. 14.6 indicates the location of the hand in each frame by a tick mark on the gesture trajectory. The density of tick marks is a measure of velocity. Sparse tick marks indicate rapid motion, whereas dense areas of tick marks indicate slow motion. The velocity profile indicates pre- and poststroke holds for two gestural strokes: one on the ESE-ward (East–Southeast) stroke from Ballast Point through the EP, and the other on the SSW (South–Southwest) stroke from Bravo Wharf ~~and~~ the EP. These holds give special salience to these sections of the gestural trajectory.

Another useful cue is the shape of the gestural trajectory. Because lines of position are, by definition, straight, gesture segments that are curved are unlikely to be meaningful representations of virtual lines of position. Correspondences between potential lines of position and the linear segments of the gesture add plausibility to some potential lines of position and not others. Again, this cue is not, by itself, sufficient to pick out the lines of position being proposed by the bearing recorder. The straightest section of the gesture trajectory does not correspond to any possible LOP.

Some parts of the gesture are performed many centimeters above the surface of the chart, whereas others are performed with the tip of the finger in contact with the chart. Real lines of position are drawn by putting a pencil in contact with the surface of the chart. Making contact with the chart seems to add perceptual salience to these parts of the gesture. The two strokes that correspond to the intended lines of position are made with the tip of the finger in contact with the surface of the chart.

Finally, one can add the cogesture speech to the representation. The bearing recorder says, “It’ll be that ‘n that.” The two occurrences of the indexical “that” are produced in synchrony with the two meaningful strokes and add to their perceptual salience. These words mediate the allocation of attention, of speaker and listener, to the gestural performance. These words imply a structure of relationship among the elements of the multimodal system (something will be composed of two parts), but the identity of the elements and the nature of their relationship is not in the words alone; it is in the interpretation of the environmentally coupled gesture.

The combined contributions of these cues unambiguously pick out two gestural strokes as representations of proposed lines of position. These are not the straightest strokes in the gestural trajectory, nor are they exactly aligned with landmark or estimated position, but the combination of cues marks them as unambiguously meaningful. The meanings of the motions that constitute the gesture are established by their relations to the other elements of this complex act of meaning making.

No one knows in what order or how these cues are perceived, processed, or combined. This is precisely the problem indicated by Levinson under the heading of the “binding problem.” (Levinson this volume). Multimodal signal streams require the linking of elements that belong to one another across time and modality. Although all of the cues have discernable physical properties, it is not the signals themselves that make the cues relevant. It is the meaning that the overall performance has as part of the understood project at hand.

### *Thinking with Brain, Body, and Culturally Constructed World*

The cultural practice of gesturing in meaningfully interpreted space brings the objects of interest, potential lines of position, into existence. This is an example in which high-level cognition is enacted in the motion of the body in shared culturally meaningful space. It is also likely that this cultural practice takes advantage of some very general properties of brain organization.

The distributed cognition perspective makes the boundary around the person permeable and leads to a natural curiosity about the relations of the other cognitive structures to activity in the brain. Unfortunately, little is known about how the brain accomplishes high-level cognition. In recent years, a large number of brain imaging studies have hinted at the promise of being able to localize some kinds of processes in regions of space (fMRI) or time (ERP), but the actual mechanisms remain obscure. For example, retinotopic maps in the visual system exist in low-level visual areas, but at higher levels the patterns of activation turn into something other than topological variants of retinal representation, and the significance of spatial patterns of activity becomes increasingly difficult to interpret. However, even without knowing the details of the processes, it is possible to say some things confidently about brain and cognition from a distributed cognition perspective.

The simple acts of seeing the landmarks and the ship’s estimated position on the chart bring visual processes into coordination with

structure in the chart and with memories for the depiction of the landmarks. This is already a complex process because the memory may be recall of specific depictions of known landmarks and/or recognition of landmarks through the interpretation of the graphical conventions used in cartography. In either case, these marks on the chart are recognized as depictions of landmarks and the previously plotted estimated position. The visual and somatosensory systems produce many representations of the location of the points of interest and the spatial relations among them. There are certainly retina-centered representations, but probably also head-centered and body-centered representations as well (personal communication, M. Sereno, October 5, 2003). Each representational system may encode multiple features such as location, direction of motion, and velocity. High-level conceptual and visual constraints on what a LOP can look like and where it can occur on a chart support the imagination of possible lines of position. This may be coordinated with eye movements tracing the LOP or saccading between the depiction of the estimated position and the depictions of the landmarks. Thus, simply seeing the chart as a meaningful space is already a complex cognitive activity.

So why gesture? By superimposing gesture on the meaningfully interpreted chart surface a navigator adds representations of motion to the visual system and representations of the trajectories of motion of the hand and fingers to the somatosensory system. At present, no one knows exactly how these representations work, but imaging studies show that there are from ten to fifteen parietal areas carrying coordinated representations of space and motion in space (personal communication, M. Sereno, October 5, 2003). The hands, guided by conceptually meaningful visual and motor representations, act in the world thereby producing new richer more complex and more integrated brain representations. By acting and monitoring one's own action at the same time one uses brain processes to guide activities that entrain more brain processes. This is a self-organizing process that is located in the brain-body-world system.

Reasoning about the angles of intersection of the LOPs requires stable representations of the LOPs. The robustness of the high-level cognition depends on the way this activity coordinates a large number of related representations, some in the environment of action, some in the body, and some in the brain. The cultural practices take advantage of the way the brain works to bring into existence multiple representations that together are more stable than any single representation alone.

The practice the navigators engage in is located in a complex cognitive ecology. The practice of gesturing to imagine lines of position brings into coordination many elements in a rich web of constraints that includes the technological tools of the job, the social relationships and division of labor among the people, the functional organization of the brain, and the culturally shaped ways of using the body. The high-level cognitive accomplishment, choosing appropriate landmarks, depends on all of these things. Each element of the system makes sense in the context of its relations to the other elements. This tight web of interrelationships is typical of real-world cognitive ecologies. In such systems the correct unit of analysis is not one brain or even one semiotic modality, such as speech or gesture taken in isolation, but the entire system. The meaning of a complex emerges from the interactions among the modalities that include the body as well as material objects present in the environment. The effects of these interactions are generally not simply additive. Such a meaning complex may be built up incrementally or produced more or less whole, depending on the nature of the components and the relations among them (see Alač and Hutchins 2004; Goodwin 1994, this volume; Hutchins and Palen 1997).

Navigation is a special domain of activity, and this sometimes gives rise to concerns about the generalization of the findings made here. Fortunately, navigation does not involve cognitive processes that are alien to everyday life. Rather, what is special about this setting is how well it supports enacted reasoning. The generalization of results must be tied to the distribution of the mode of thinking, not to the characteristics of the setting. We now have ample evidence that enacted reasoning is surely a very widespread phenomenon. Goodwin (this volume) highlights the importance of the meaningfully interpreted material world when he says that environmentally coupled gesture is pervasive. Others in this volume who describe the central involvement of meaningfully construed material environment include Byrne, Clark, Enfield, Gaskins, Goldin-Meadow, Hanks, Keating, Levinson, Liszkowski, and Tomasello. The observations reported by these authors span species, cultures, and levels of development. It is therefore likely that embodied reasoning is a very old and widespread cognitive process.

### *Implications: Being Human*

Although researchers are increasingly attending to interactions in which the physical and social environments for action play an important role, that role is still not clearly conceptualized. In the example presented



above, we saw that the social distribution of cognitive labor increases the variability in the choice space. The conversational practice of taking turns suggesting and evaluating options creates a cognitive system that is likely to explore a wider range of alternatives than would be explored by any navigator alone. Our folk theories assume that thought precedes action. I have tried to show that in some activity settings, acting in the world *is* thinking (see also Alač and Hutchins 2004). Finally, processes of cultural evolution can produce activity settings in which simple courses of action can produce powerful cognitive processes.

With these observations, I offer a sketch of an image of interaction as a complex dynamic system. Typical human–human interactions are composed of many elements, the meanings of which emerge from the network of relations among the elements. For example, the representations of the provisional imagined LOPs are emergent properties of the complex activity system. They cannot be partialled out as being  $x$  percent in the brain,  $y$  percent in the body, and  $z$  percent in the world. Like the components of a position fix, the parts of a meaningful human interaction only mean what they mean by virtue of their roles in the whole culturally understood activity.

### *Implications: Becoming Human*

All serious cognitive scientists acknowledge the importance of symbolic processes in human cognition. But where, when, and how are symbols involved in human cognition? As noted in the discussion of the nature of interaction above, much more work needs to be done to document the distribution of cognitive strategies across space, culture, and context. Although internal symbol processes must be inferred from observable behavior, the use of external symbols is quite apparent. And this provides the basis for some speculations about symbolic processes.

In a seminal work, Rumelhart et al. (1986) argue that individual humans are good at three sorts of activity: (1) recognizing patterns, (2) manipulating the physical world, and (3) imagining simple dynamical processes. They describe how these processes could be invoked by a person doing place-value multiplication with paper and pencil.

Each cycle of this operation involves first creating a representation through manipulation of the environment, then a processing of the (actual physical) representation by means of our well-tuned perceptual apparatus leading to further modification of this representation. By doing this we reduce a very abstract conceptual problem to a series of operations that are very concrete and at which we can become very



good. . . . This is real symbol processing and, we are beginning to think, the primary symbol processing that we are able to do. [Rumelhart et al. 1986:46]

In the example of the navigators enacting lines of positions, we see that manipulating the world and imagining the dynamics of simple worlds happen together. Environmentally coupled gestures allow the navigators to use the motion of their bodies to imagine prospective lines of position. In doing this, they are reasoning about properties of the relations among the enacted LOPs. The representations of interest here do not exist until they are enacted in the world of action. They come into being as external representations created in the complex interactions of the navigators with each other and the technology of the job. Once these representations have been created in ephemeral external form, the consequent multiple coordinated internal images of them have the persistence needed to support reasoning about the angular relations among them.

Once LOPs have been experienced as external representations, they can be imagined. A navigator could even, perhaps, imagine gesturing, thereby creating imagined enacted prospective LOPs, although one suspects that the results of such imagining would not be as stable or persistent as the results of actually making the gestures. As Rumelhart et al. note, “Not only can we manipulate the physical environment and then process it, we can also learn to internalize the representations we create, ‘imagine’ them, and then process these imagined representations—just as if they were external” (1986:47). This story does not explain how external representations arise, but it does claim that once external representations arise, there is a possibility of those representations being imagined by a person, and the person imagining transformation of those representations.

The argument above assumes that symbols could arise in interaction before they arise internally. Is such a thing possible? When cognition takes place in the interaction of the mind with the surrounding environment, there is a new place to look for the origins of cognitive processes and structure. This is important because so many theories of the origins of human cognitive capacities go wrong by positing special processes, modules, or evolutionary miracles that seem necessary to construct a plausible story for the development of cognitive capabilities entirely inside isolated individual brains. The origins of features of language are good examples of this. But computer simulation studies have shown that communities of agents in interaction can develop shared lexicons

(Hutchins and Hazlehurst 1995) and shared propositional structure (Hazlehurst and Hutchins 1998; Hutchins and Hazlehurst 2002).

Thinking about the roots of sociality and cognition it is a common practice to project an image of activity into the past and imagine what functional properties evolution could select for to produce a more advantageous activity. When social interaction is our target, what sort of image of interaction shall we project into the past? Projecting the image of complex, multimodal, environmentally coupled interaction into the past illuminates new possibilities for development. Change can take place anywhere in the complex interaction system. This means that one need not imagine that all mechanisms of change are lodged inside individual organisms. Just as the image of complex multimodal environmentally coupled interaction gives us a new place to look for the sources of organization of ongoing behavior; it also gives a new place to look for the developmental changes across phylogenetic time.

A very similar argument is made in contemporary evolutionary biology. Oyama (2000) argues that the system that evolves is not the genome, but the phenotype in context (see also Turner 2000). The central dogma of evolutionary biology is that all important change resides in the genome. But the system that evolves is a wider system of organism and environment in interaction. Similarly, when thinking about cognition, it is a mistake to focus narrowly on hypothesized functional adaptations of the brain. It is commonly assumed that genetic adaptations must produce a brain that is capable of the hypothesized new functional abilities. What evolves, however, is not the brain alone, but the system of brains, bodies, and shared environments in interaction. Cultural practices are as much a part of the story of cognitive evolution as are changes in brain structure. This means that important milestones in cognitive evolution could, in principle, have been achieved without any particular genetic adaptation being associated with them. A change in physical environment, for example, could lead to changes in interactive processes that could give rise to a new cognitive ability in the interaction system. This goes even for critical milestones such as the advent of symbolic representations. Once a new functional capacity arises in the interaction system, it creates new opportunities for change in the genome. This argument does not deny the role of genetic change, it only points out that the genome is but one of many elements of a complex adaptive system.

This is not to say that thinking and imagining never happen in the absence of a material world, for clearly they do. But it does say that such

processes are different in nature than thinking with the world, that they are derived from (transformations of) processes that do involve action with the world, and they generally appear later developmentally (both ontogenetically and phylogenetically) than thinking with the world. The last point is a key component of Vygotsky's (1986) theory of the social origins of mind. The kinds of thinking that has been the focus of cognitive science and psychology is likely a relatively recent add-on to a more fundamental, but, as yet, poorly understood mode of thinking with the world. No one knows the relative frequencies of thinking in these different modes or thinking across behavioral settings. And we know even less about the distribution of the ways of thinking that lie along the continuum between completely mental activity and thinking that is inextricably bound up with action in the world.

### *Understanding Interaction*

Human minds did not evolve in isolation, each wrapped tightly in a thick skull and thereby insulated from the complexities of the body and the world. We know that the brain takes advantage of minute details of the body and the body's interaction with the physical environment (Clark 2001; Quartz and Sejnowski 2002). Similarly, mind will have evolved, not in isolation from the material and social world, but in ways that weave its activity inextricably into the details of those worlds (Tomasello 2001).

If distributed cognition presents us with a world in which everything is seemingly connected to everything else, does not studying cognition become impossible? I think it certainly becomes more difficult. Understanding complex real-world interactions is more difficult than understanding systems of simple linear relations. However, in some ways, more complex problems can be easier to solve than what seem to be simpler problems. If the nature of the problem is to constrain behavior, a system of multiple interacting subsystems can provide a solution more easily than trying to get all of the constraints out of a single subsystem. For example, it is easier to account for the organization of the visual system if one recognizes that it develops in concert with the auditory system than it is to account for the organization of either system in isolation (de Sa and Ballard 1998). Such findings are part of a wider shift in the cognitive sciences is toward an increasing appreciation for rich interactions among systems at all levels of organization. People in normal interaction are in the business of creating and interpreting rich multimodal meaning complexes. Here again, sometimes solving

what looks like a more complex problem is easier than solving what looks like a simpler problem. It is easier to work out the significance of complex multiply constrained acts of meaning than it is to determine the meanings of the individual components as isolated systems. It is easier to establish a meaning for words embedded with gestures that are performed in coordination with a meaningful shared world than it is to establish meanings for words as isolated symbols.

Thus, when we approach the more complex objects of scientific scrutiny demanded by distributed cognition theory, it is not the case that explanations will necessarily be more difficult to create. They may be somewhat more complex than easy linear and modular stories, but in some cases, the explanations come naturally as side effects or by products of general principles. For example the development of a shared lexicon mentioned above.

## Conclusion

By softening the traditional disciplinary boundaries the distributed cognition perspective focuses on a new unit of analysis that encloses a complex set of interactions among brain, body, and culturally constructed world. Careful attention to the microstructure of interaction from the distributed cognition perspective leads to ~~are~~ a reconceptualization of the individual–environment relationship and suggests that this newly conceived relation has important implications for the way we confront many sorts of cognitive and anthropological problems. In particular, it provides a new place to look for mechanisms that shape both the ontogenetic and the phylogenetic development of sociality.

## Acknowledgement

This work was funded by a grant from the Santa Fe Institute's program on robustness in natural and social systems, which is supported by the McDonnell Foundation. Alisa Durán transcribed the data and suggested many elements of the analysis presented here.

## References

- Alač, M., and E. Hutchins. 2004. I see what you are saying: Action as cognition in fMRI brain mapping practice. *Journal of Cognition and Culture* 4(3–4):629–661.

- Bartlett, F. 1995[1932]. *Remembering: A study in experimental and social psychology*, 2nd edition. Cambridge: Cambridge University Press.
- Bateson, G. 1972. *Steps to an ecology of mind*. New York: Ballantine Books.
- Braitenberg, V. 1984. *Vehicles: Experiments in synthetic psychology*. Cambridge, MA: MIT Press.
- Bruner, J., R. Olver, P. Greenfield, et al. 1966. *Studies in cognitive growth: A collaboration at the center for cognitive studies*. New York: John Wiley and Sons.
- Clark, A. 2001. *Mindware: An introduction to the philosophy of cognitive science*. Oxford: Oxford University Press.
- Cole, M., and M. Griffin. 1980. Cultural Amplifiers Reconsidered. In *The social foundations of language and thought. Essays in honor of Jerome Bruner*, edited by D. Olson, 343–364. New York: Norton.
- de Sa, V., and D. H. Ballard. 1998. Category learning through multimodality sensing. *Neural Computation* 10(5):1097–1117.
- Douglas, M. 1986. *How institutions think*. Syracuse, NY: Syracuse University Press.
- Enfield, N. J. 2005. The body as cognitive artifact in kinship representations: Hand gesture diagrams by speakers of Lao. *Current Anthropology* 46(1):51–73.
- Goodwin, C. 1994. Professional vision. *American Anthropologist* 96(3): 606–633.
- Halbwachs, M. 1925. *Les cadres sociaux de la memoire*. Paris: Albin Michel.
- Hazlehurst, B., and E. Hutchins. 1998. The emergence of propositions from the coordination of talk and action in a shared world. *Language and Cognitive Process* 13(3):373–424.
- Henty, S. 1999. A computational simulation of jury decision making. Honors thesis, Department of Cognitive Science, University of California, San Diego.
- Hollan, J. D., E. Hutchins, and D. Kirsh. 2001. Distributed cognition: A new foundation for human-computer interaction research. *ACM Transactions on Human-Computer Interaction: Special Issue on Human-Computer Interaction in the New Millennium* 7(2):174–196.
- Hutchins, E. 1995a. *Cognition in the Wild*. Cambridge, MA: MIT Press.
- . 1995b. How a cockpit remembers its speeds. *Cognitive Science* 19:265–288.
- . 1999. Cognitive artifacts. In *The MIT Encyclopedia of the Cognitive Sciences*, edited by R. A. Wilson and F. C. Keil, 126–128. Cambridge, MA: MIT Press.

- . 2000. The cognitive consequences of patterns of information flow. *Intellectica* 1(30):53–74.
- Hutchins, E., and B. Hazlehurst. 1991. Learning in the cultural process. In *Artificial life 2*, Santa Fe Institute studies in the sciences of complexity series, edited by C. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen, 689–706. Santa Fe: Santa Fe Institute.
- . 1995. How to invent a lexicon: The emergence of shared form-meaning mappings in interaction. In *Social intelligence and interaction*, edited by E. Goody, 53–67. Cambridge: Cambridge University Press.
- . 2002. Auto-organization and emergence of shared language structure. In *Simulating the evolution of language*, edited by A. Cangelosi and D. Parisi, 279–306. London: Springer Verlag.
- Hutchins, E., and T. Klausen. 1996. Distributed cognition in an airline cockpit. In *Cognition and communication at work*, edited by Y. Engeström and D. Middleton, 15–34. New York: Cambridge University Press.
- Hutchins, E., and L. Palen. 1997. Constructing meaning from space, gesture, and speech. In *Discourse, tools, and reasoning: Essays on situated cognition*, edited by L. B. Resnick, R. Saljo, C. Pontecorvo, and B. Burge, 23–40. Heidelberg: Springer Verlag.
- Luria, A. M. 1966. *Higher cortical functions in man*. New York: Basic Books.
- McNeill, D. 1992. *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.
- Middleton, D., and D. Edwards (eds.). 1990. *Collective remembering*. London: Sage.
- Newell, A., and H. Simon. 1972. *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Norman, D. 1994. *Things that make us smart: Defending human attributes in the age of the machine*. Boston, MA: Addison-Wesley.
- Oyama, S. 2000. **Evolution's eye: A systems view of the biology-culture divide**. Durham, NC: Duke University Press.
- Quartz, S., and T. Sejnowski. 2002. *Liars, lovers, and heroes: What the new brain science reveals about how we become who we are*. New York: Morrow.
- Roberts, J. M. 1964. The self-management of cultures. In *Explorations in cultural anthropology*, edited by W. H. Goodenough, 433–454. New York: McGraw-Hill.
- Rumelhart, D., P. Smolensky, J. McClelland, and G. Hinton. 1986. Schemata and sequential processes in PDP models. In *Parallel distributed processing*, vol. 2, edited by J. McClelland and D. Rumelhart, 7–57. Cambridge, MA: MIT Press.

- Surowiecki, J. 2004. *The wisdom of crowds*. New York: Doubleday.
- Tomasello, M. 2001. *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.
- Turner, J. 2000. *The extended organism*. Cambridge, MA: Harvard University Press.
- Vygotsky, L. S. 1986. *Thought and language*. Cambridge, MA: MIT Press.