Imagining the Cognitive Life of Things

Edwin Hutchins (ehutchins@ucsd.edu)
Department of Cognitive Science
University of California San Diego

Abstract

Human interaction is complex. An embodied perspective on interaction shows that it is richly multimodal. The complexity of the interaction system allows for a surprising variety of emergent cognitive effects. This paper attempts to place the cognitive life of things in the context of rich multimodal interactions.

Part I: The cognitive life of things is manifest in the ways people jointly engage things with their bodies and their words
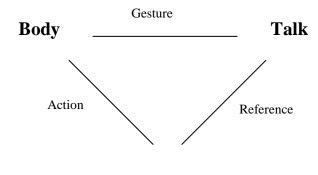
In the last chapter of *Cognition in the Wild* (Hutchins, 1995), I argue that cognitive science made a fundamental category error when it mistook the properties of a person in interaction with a social and material world for the cognitive properties of whatever is inside the person. One enduring problem with this claim is that it demands a description of how cognitive properties arise from the interaction of person with social and material world. *Cognition in the Wild* provides a profoundly incomplete answer to this question. In the years since it's publication, *Cognition in the Wild* has been criticized for saying so little about the people in the navigation setting. It describes the tools of the trade, and the historical development of the tools. It describes social processes and the cognitive properties of those social processes, but it says almost nothing about the embodied practices of the navigators as flesh-and-blood people. For the most part, the cognitive processes described in *Cognition in the Wild*, and in other treatments of distributed cognition, are presented without reference to the role of the body in thinking. That is, in spite of the fact that distributed cognition claims that the interaction of people with things is a central phenomenon of cognition, the approach has remained oddly disembodied.

The multimodality of verbal behavior

For the past three months, I have been teaching my undergraduate course on the methods of cognitive ethnography. Late in the quarter we transcribe and analyze spoken discourse. When dealing with language materials, the students learn that everyday language is complex. The transcription methods we use (borrowed from Gail Jefferson (1984)) capture the words that are spoken, but also focus attention on super-segmental features of timing, prosody, stress patterns; the aspects of language that are most affected by the fact that speech is produced by a body in action. We seem to forget sometimes that speaking is an embodied activity. After transcribing at the word and morpheme level, I try to help the students see the morphological, lexical, syntactic, prosodic, and pragmatic organization of talk. Then I have students identify the cultural models that account for the discourse level organization of the talk. What underlying assumptions hold the parts of a discourse together? By the time they have done this, it is apparent that verbal behavior is richly multimodal in the sense that it integrates many layers of structure, each consisting of a distinct set of patterns, and all coordinated with one another by a very complex set of relations.

The richness of embodied interaction

For the final project in the course, I have the students collect, transcribe, and analyze video recording of real-world interaction (often in a work setting). Students are often astonished to discover how much more complex video records of interaction are than audio records of spoken narratives. In addition to talk, video recordings include the bodies of the participants, and the world that the interactants co-inhabit. I show my students a diagram (figure 1) to illustrate the increased richness of the video data compared to audio.

Figure 1: A multi-modal interaction system

Just as there are multiple complex patterns in verbal behavior, there are also multiple complex patterns in the deployment of the body (hand gestures, of course, but also other limbs, body orientation, head position and motion, eye-gaze, facial expression, and more) and often very complex patterns (arrays of things) in the world that the interactants share. Furthermore, with the additional elements come new relationships among the elements. Some of these relationships have already spawned entire

fields of inquiry: co-speech gesture, environmentally-coupled gesture, and reference for example. This growing appreciation for the cognitive importance of the details of body-world interactions owes a lot to the availability of inexpensive digital video. Frame-by-frame examination of interactions of people with one another and with their environments for thinking makes it possible to see previously hidden detail.

The diagram and this discussion are gross oversimplifications, of course. The uses of the body are also richly multimodal (as we will see below), as are interactions with the world. Thus each mode depicted in the multi-modal system diagram is itself a multi-modal system.

Situating embodied interaction in the social world
The diagram in figure 1 is also an oversimplification because it leaves out the embodied nature of social interactions. For the most part, when people engage things cognitively, they do so in social contexts. Consider the added complexity that comes when we consider interaction systems in which multiple actors are engaged in joint activity.
Addressing the embodied nature of social interaction requires a change to the diagram. Rather than a single body and a single stream of talk, we now have two bodies, and two streams of talk, and a single shared world as shown in figure 2. I call this diagram the "square-cut gem" because of its inverted tetrahedron shape.

Body ——————— **Body**
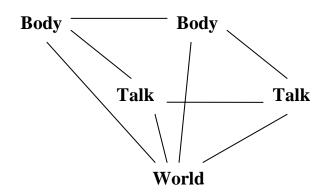
Talk ———— **Talk**

World

Figure 2: A square-cut gem of interaction

Now, instead of adding new elements, we are only adding new relations. The new relations hold between body and body, between talk and talk and between the body of one and the talk of another. The many cognitive properties of systems of socially distributed cognition also enter at this level.

There are six kinds of edges in the square-cut gem of interaction. The edges are loci of coordination and each is characterized by a distinct set of coordination types. In addition to edges, there are facets on the gem, which correspond to more complex systems of multimodal coordination. We are already familiar with the B/T/W facet for an individual. We have now added three facets: B/B/W, T/T/W, and a facet that would be called the "table" on a gem, B/B/T/T. A typical conversation between two people who are not handling artifacts might be thought to take place at the table.

The T/T/W facet is the proper location of discourse analysis, although we have already begun our discussion of it above. The coordination processes that give rise to turn taking, sequential contingency of meaning, projection, and intersubjective reference and anaphora belong here. B/B/W comprises relations of coordination (entrainment of bodily motion as in synchronous applause), cooperation (passing on the street, e.g.) and collaboration (lifting something heavy together). Also the establishment and maintenance of shared attention, as well as the recipient design of gestures are forms of coordination on this facet. Talk1/Body2 can be seen in conventional greetings. Person 2 waves, or nods, person 1 says "Hey!" . Also in the microstructure of on-going interaction in which a speaker shapes what is said based on the on-going non-verbal reactions of the interlocutor (Hindmarsh and Heath, 2000).

The most interesting interactions take place in a system that involves the entire gem of interaction. Figure 3 shows two pilots working together in a complex setting: the flight deck of a Boeing 747-400. They are jointly planning an approach at their destination airport. A quick look here highlights some of kinds of coordination that arise when embodied interaction is situated in the social realm.

Figure 3: Pilots jointly imagine the future

The system of body/world coordination is rich. In addition to the pilot on the left locating and opening to the relevant page in the route manual, the two pilots are engaged in mutual recipient design in talk and in gesture. Each is shaping his own verbal and non-verbal behavior to suit the needs of the other. Relations of talk to the world demonstrate constraints in both directions. The spatial organization of a thing (airport description page) is being used to create sequential temporal structure in the activity by adopting a spatial "flow" of attention across the page. Reciprocally, talk constrains the engagement of the thing by making deictic reference to the items on the page. Talk and

body are coordinated as co-speech gesture is used to highlight items referred to and to guide own and other's attention. Reciprocally, co-gesture speech provides clues on how to identify the meaningful segments of a complex fluid gestural movement. Both pilots use their body orientations and postures (construed within the constraints of the work stations) to demonstrate shared attention and interest. The relations among the two speaking roles are complexly organized by social and professional roles and an intersubjectively shared understanding of the nature of the task. Using these means and others, the pilots jointly imagine the future of their airplane from their current location, 150 miles from their destination, all the way to the arrival gate. The human interaction system is composed of a very complex network of coordination among myriad dynamic behavioral patterns. I propose to use this vision of interaction to further explore the cognitive life of things.

The power of embodied interaction with culturally constructed settings.

The power of cognitive artifacts to transform human thought is widely recognized. But of what importance are the details of the bodily interaction with a cognitive artifact in producing its cognitive power? The study of cognitive artifacts has not attended closely to the nature of the on-going activities in which people think in interactions with things. This important topic has dropped into the interstices between academic disciplines. Of course, ergonomics considers the fit between the properties of bodies and the manipulation of tools. But, we do not yet have a tradition of thought that appreciates the profoundly situated character of embodied thinking in complex culturally organized settings.

In my work on material anchors for conceptual blends (Hutchins, 2005), I tried to show how in some kinds of activities, physical relations become proxies for conceptual relations. In the linguistic phenomenon known as fictive motion, a conceptual relationship is mapped onto a real or imagined space to produce a trajectory to guide the allocation of attention. In the well-known memory technique called the "method of loci," and in the techniques that Kirsh (1995) groups under the heading "the intelligent use of space," an imagined trajectory is mapped onto real things in space, thus producing sequential relations among the things. My analysis of material anchors for conceptual blends was, however, like most other work on cognitive artifacts, disembodied.

The body did come into play at one point in a way that I can now see as suggestive. Some Japanese school kids learn a technique for mapping the names of months and days of the week onto their left hand in a way that allows the hand to be used as a date calculator[1]. The mapping of names of months onto the segments of the first three fingers of the left hand is learned as a sequential pattern of touching the tip of

the thumb to the segments. It is clear that the motor memory for the touch pattern is learned before the visual representation of the array of month names that can be produced by running the touch pattern in coordination with the standard list of month names. In this case, the body plays a key role in the construction of a complex material anchor. The hand is opportunistically appropriated as a "thing" possessing a set of regions and a learned motor routine controls the assignment of conceptual content to the regions. In interaction with things, embodied processes can become "somatic" anchors for conceptual blends; they are internal proxies for abstract relations among the things with which one interacts. I will return to this point in an extended example below.

Let us return now to the rich systems of interaction that are visible in video recordings. An emerging field of inquiry can be seen in the work of a number of researchers who focus on the details of the creation and manipulation of material representations. In domains as varied as scientific research (Goodwin, 1994; Alač, 2003), architectural practice (Murphy, 2004), student learning (Singer and Goldin-Meadow, 2005, Alač and Hutchins, 2005), airline piloting (Hutchins and Palen, 1997), telecommunications troubleshooting (Hindmarsh and Heath, 2000), and surgery (Goodwin, n.d.), researchers are showing how interactions between the body and cultural artifacts constitute an important form of thinking. These interactions are not taken as "indications" of invisible mental processes, rather they are taken as the thinking processes themselves. In this brief paper I will try to add to this project. I will illustrate how processes of imagination, inference, error detection, and even "Aha!" insight can arise in cultural practices that bring the body into interaction with the cultural environment in particular ways.

Using the body to imagine the dynamics of things

The notion of somatic anchors for conceptual blends adds a new dimension to the now familiar phenomena of environmentally-coupled gestures (Goodwin, Hutchins and Palen, Becvar, Alac many more). Let me give just two brief examples here in which scientists engage static graphical representations with their moving bodies in order to imagine the complex dynamics of entities that cannot be observed directly. The first example comes from the work of my student, Amaya Becvar (Becvar, Hollan, Hutchins (in press a, and in press b)). Figure 4 shows the principal investigator of a molecular biology laboratory. The problem at hand is to imagine how the shape of the thrombin molecule changes when another molecule, thrombomodulin, attaches to it. She has just aligned her left hand with a ribbon diagram of a thrombin molecule that is depicted on an overhead transparency lying on the platen of an overhead projector. She lifts her hand from the transparency as if, by virtue of touching her fingers to the loops in the ribbon diagram, her hand has become the molecule. She then says, "Our theory is that the molecule moves either like this or like this."

---

[1] See Nakahara (1996) for details.

Coincident with each of the two deictic utterances of "this" she produces a motion with her hand. In one the finger tips come together like the mouth of a purse. In the other, the fingers rotate around an imaginary axis emerging from the palm of the hand. In this way, she imagines two dynamic modes for the loops of the thrombin molecule.

Figure 4: The thrombin hand

The second example comes from the work of another of my students, Morana Alač (Alač, 2003). In figure 5, we see a brain imaging researcher positioning his hands in front of a computer display showing a functional image of a brain. His hands are cupped and they enclose a volume that is about that of a three-dimensional human brain. He says, "You can see it's sort of …" During the pause following the word "of" he moves his left hand up and his right hand down. Subsequent talk and action show that he is referring to the shearing of the image – an artifact introduced by the software that computes the image. In this way, he uses his body to imagine a process, "shearing," that is held responsible for the observed asymmetry of the image. This is an interesting bit of scientific reasoning because the imagined process is both very useful and completely fictional.

Figure 5: The sheared brain

In these two examples, the researchers use the motion of their own bodies in coordination with static representations to imagine dynamic properties of entities that cannot be seen or sensed directly. These bodily motions are the medium in which the researchers reason about the entities that are depicted in the static representations.

Imagining virtual worlds
Similar work by Murphy (2004) shows how architects imagine the dynamics of people moving in buildings that do not yet exist. He describes environmentally-coupled gestures that represent imagined motion along paths in the as-yet-unbuilt building. He gives these gestures the label, "action in the subjunctive mood." One important insight here is that these representations are tentative in a way that the lines that appear on the building drawings are not. They suggest that representations can have different degrees of commitment to the entities that are represented. Some representations make strong commitments, while others are more tentative. Some are tangible and permanent (lines in ink on a building drawing) while others are less permanent (pencil markings), while still others are ephemeral (gestures). I will argue below that understanding the cognitive life of things requires that we understand the full range of commitment of our embodied engagements with things.

Relations among representational contents in multimodal representations
We need to consider one more recently described phenomenon before taking on an extended example. Susan Goldin-Meadow (Singer and Goldin-Meadow, 2005) describes a "mismatch" between talk and environmentally coupled gesture in interaction as an index of readiness to learn. Mismatch may be the wrong word here, but the insight that different media in a multimodal interaction may carry different information is key. It is possible to distinguish cases in which gesture and talk carry overlapping or congruent representations from those in which gesture and talk carry different, but complementary, representations. The latter cases are labeled "mismatch" by Goldin-Meadow, and they are taken to indicate that a child is on the verge of conceptual breakthrough. Goldin-Meadow is describing emergent properties of a multimodal system of representations (although she does not use the term "emergent"). In the example below, I will argue that the embodied imagining that is made possible by complementary representations in the multi-modal body/world system can give rise to an "Aha!" insight. The additional twist here is that the complementary representations are not carried in talk and gesture, but are both are contained in the multiple modalities of the relations of body to the world of things. A new insight arises as an emergent property of the juxtaposition of complementary content in distinct modalities.

Part II: An example of thinking with things in ship navigation

Until recently, ship navigation was performed on paper charts using manual plotting tools. The data on which this analysis is based were originally collected more than 20 years ago on the bridge of a US Navy ship when these practices were still common. In order to fix the position of a ship, navigators measure the bearing from the ship to at least three landmarks. When plotted on a chart, the bearing of a landmark from the ship becomes a line of position (LOP), that is, it is a line on which the ship must be located. Plotting a LOP involves setting the measured bearing on a

protractor scale on a plotting tool (called the hoey) and then placing the hoey on the chart so that the protractor arm passes through the depiction of the landmark on the chart and the base of the protractor scale is aligned with the directional frame of the chart. Once the plotting tool is correctly placed, the navigator uses a pencil to draw a line on the chart in the vicinity of the projected position of the ship. Two intersecting lines of position determine, or "fix", the position of the ship. Navigators usually try to plot three lines of position, because the intersection of three LOPs forms a triangle. A small fix triangle indicates that the position fixing information is good. A large triangle indicates problems somewhere in the chain of representations that lead to the fix triangle. In general, the navigator's level of anxiety is proportional to the size of the fix triangle.

In order to produce a useful fix triangle, the spatial relationships among the ship and the three landmarks must be such that the LOPs intersect at steep angles. This means that navigators must anticipate where the ship will be when the next set of position fixing observations is made and must choose a set of three landmarks with appropriate relations for the fix.

A gradient of tangibility in the representations of LOPs.

The video data contain records of activities that create many sorts of representations of LOPs. Each system is an integrated whole that brings representations in multiple modalities into coordination. Each has distinct cognitive properties. The observed systems include the following:

1. Situated seeing of the chart. Navigators look at the location of the estimated next position of the ship and name sets of landmarks that may produce the appropriate angles between LOPs. In this case, the representations of the LOPs are produced in visual imagination superimposed on the chart surface.
2. Enacting LOPs with finger motions in the air above the chart. In these cases, the visual imagination of the LOPs is augmented by gestures made approximately 0.5 meters above the chart surface. This practice adds gesture to the system and integrates motor, visual, and proprioceptive representations.
3. Tracing LOPs with finger motions on or very near the surface of the chart. In these cases, the visual imagination of the LOPs is augmented by gestures made in close coordination with the visual structure of the chart. Making contact with the chart surface while gesturing integrates tactile representations into the system. This practice is especially important in episodes of joint reasoning in which

navigators choose sets of landmarks to use for future fixes[2].
4. Positioning the straight edge of the plotting protractor (hoey) in the approximate location of LOPs and manipulating the position of the straight edge to imagine variations on previously plotted LOPs. In this case, the navigator uses the plotting tool's straight edge to temporarily imagine/create/locate approximate LOPs. Manipulation of the hoey integrates motor, visual, proprioceptive and tactile representations. Examining the placed tool adds stable visual representations to the system.
5. Plotting LOPs with the hoey. In these cases, the hoey is exactly aligned with the landmark and the directional frame of the chart, and a pencil line is drawn on the chart. This integrates visual, motor, proprioceptive, and tactile representations and also produces a durable representation of the LOP on the surface of the chart.

Across these five cases, one sees increasing tangibility and accuracy of the LOP representations, increasing involvement of the body and the culturally organized setting, and increasing cognitive and physical cost. It is worth noting that the fifth practice is the only one that is formally part of the position plotting "task." As such, it is the only one that would be likely to be examined in a study of the ergonomics of position plotting. But the others are essential to understanding how navigators actually accomplish their work. I will argue below that it is necessary to understand the entire range of practices described here in order to understand how embodied interactions with cultural artifacts can support abstraction.

In the middle region of this gradient (cases 2-4) are actions that Murphy (2004) has called "action in the subjunctive mood." These are "as-if" actions or "may it be thus" actions. These actions produce ephemeral representations of potential, but not yet realized, states of affairs or processes. Navigators reason about the relationships among the virtual LOPs that are created by these gestures. These are thus clear examples of thinking with the body in interaction with a cultural artifact. Notice that the boundaries of gesture and instrumental action become blurred in this continuum.

I will now focus briefly on the fourth class, positioning the hoey arm and manipulating the position of the arm to imagine LOPs. This class of representational system was observed in an event that was a crisis for a navigation team. While entering a narrow navigation channel, a ship suffered the failure of its main gyrocompass. Upon losing the gyrocompass, the navigation crew could no longer simply read the true bearing of a given landmark and plot that

---

[2] A more detailed discussion of this system appears in Hutchins (in press).

bearing. Rather, they were then required to compute the true bearing by adding the corrected magnetic ship's head to the relative bearing of the landmark. The correct equation is: true bearing equals compass heading plus deviation plus magnetic variation, plus the relative bearing of the landmark. (TB = C + D + V + RB). This perturbation disrupted the ability of the crew to plot accurate positions for the ship. The crew explored various computational variations of TB = C +V +RB for 38 lines of position. Then they discovered that a key term, deviation, was missing from their computations. After reconfiguring their work to include the deviation term, the team gradually regained the functional ability to plot accurate positions.

The discovery of a missing term

How can the discovery that this term was missing be explained? The discovery appeared as an "Aha!" insight. Taken in the context of the computations that the crew was doing, this discovery is, like most creative insights, mysterious. There is nothing in the pattern of computational efforts leading up to the discovery that indicates they are nearing this development. The processes that underlie the "Aha!" insight remain invisible to a computational perspective in part because that perspective represents everything in a single mono-modal (or even a-modal) system[3]. A careful examination of the way the body engages the tools in the setting, however, helps to solve the mystery of how the discovery was made, and why it happened when it did. The insight was achieved in and emerged out of the navigator's bodily engagement with the setting.

Here is a very brief account of the course of events:
Lines of position had been plotted to each of the three landmarks, but the fix triangle that was produced was unacceptably large. Such a large triangle was clear evidence of the presence of an error somewhere in the process that created it. The LOPs were then checked, and at least one possible source of error was tested with respect to each one. Then the plotter then did the following:

1. He aligned the hoey approximately for one landmark, them moved it slightly clockwise with respect to the previously plotted LOP. He then quickly shifted the hoey on the surface of the chart and aligned it approximately with the second landmark. He also adjusted this LOP slightly clockwise. In these two moves, the plotter was exploring the lines of position and using his body and the tools (chart and hoey) to imagine LOPs that would make the fix triangle smaller.
2. He spoke (self-regulatory speech) the remembered bearing to the third landmark, 120 degrees, while the hoey was still lying on the chart. He then

picked up the hoey and moved the arm in the direction of the 120 degree scale position. It is not possible to establish exactly where the hoey arm was on the scale when he stopped moving it, nor can one determine whether or which scale values he had read at that point. It is clear that he was moving the hoey arm across the scale and attending to the scale values.

3. While moving the hoey arm toward alignment with the scale at the 120 degree mark, the plotter was also imagining the location with respect to the fix triangle of the LOP that was plotted using 120 degrees. With his body he imagines a slight clockwise rotation of the LOP that would make the triangle smaller. Several representations are integrated into a single meaning complex at this moment. They include the hoey arm and protractor scale and the imagined clockwise rotation of the hoey arm on the chart. This combination of representations creates the image of a small clockwise motion of the arm with respect to the hoey base that would decrease the size of the triangle (see figure 6). In this composite image, it can also be seen that a clockwise motion corresponds to motion across the scale toward larger numbers (see figure 7). This implies a slightly larger numeric value for the bearing to that landmark.
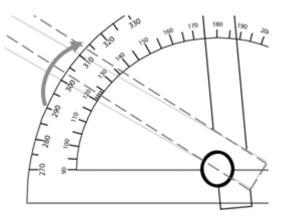

Figure 6: The superimposition of imagined clockwise rotation onto the hoey plotting tool.

---

[3] In Hutchins (1995), I provide a disembodied analysis of this event that fails to explain how the discovery of the missing term was made.
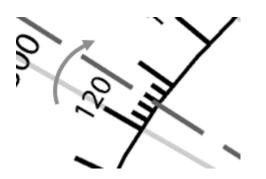
Figure 7. Seeing that clockwise rotation produces larger numbers.

4. The plotter then seems to remember that the explored actions to improve the other two LOPs was also clockwise rotation (the actions that established those relations occurred three and eight second earlier, respectively). This implies that all three of the bearings should be slightly larger numbers.

5. The final step in the process is noticing that the need to add a small amount to each bearing suggests that a term may be missing from the correction calculation. Unfortunately, the processes that produce this last step are hidden from our view. However, very conventional processes of semantic association could be invoked to explain this final step because the elements that are created by the complex of embodied representations closely match the elements of the schema for deviation. It is no longer a mysterious process.

In a traditional cognitive explanation of creative insight, one would postulate the entire discovery process in terms of interactions among unobservable internal representations. What makes such accounts mysterious is that we know so little about these internal representations. By construing the engagement of the body with the things in the working environment as a form of thinking, we can directly observe much of the setup for the insightful discovery. We can now ask: what opportunities for noticing conceptual relations are facilitated by the means of representation that are created in the engagement of the body with the culturally constituted world of action? The bodily anticipation of clockwise rotation becomes a somatic anchor for the concept of increasing bearing number value. The serendipitous combination of tool-based representations and the imagination of future bodily motions produces an image of adding a small amount to each of three bearings.

From the perspective of a formal representation of the task, the means by which the tools are manipulated by the body appear as mere implementational details. They are a source of variability in performance, but are formally irrelevant to the accomplishment of the task. It is worth pointing out, however, that this variability in task irrelevant dimensions

may be a resource for adaptive processes when routine activity is disrupted.

Part III: Understanding the cognitive life of things is essential to understanding cognitive life in general

The processes described above can be characterized in terms of some general implications of the embodied enacted view of cognition.

- In certain culturally constructed settings, bodily motion acquires meaning by virtue of its relation to the spatial structure of things. Goodwin calls this phenomenon "environmentally-coupled gesture." In some circumstances, the body itself becomes a cognitive artifact, upon which meaningful environmentally-coupled gestures can be performed (Enfield, 2005; Hutchins, 2005). In such settings, motion in space acquires conceptual meaning and reasoning can be performed by moving the body. Courses of action become trains of thought. For example, when working on the chart, movement away from the body is conceptually northward, toward the body is south, clockwise rotation is increasing measure of degrees. When actions are performed by experts in these domains, the integrations of bodily sensation with directional frame produces embodied reasoning. Navigators sometimes speak of their reasoning skills in as "thinking like a compass." I believe this could be better described as "having bodily sensation like a compass."
- Practitioners who live and work in complex culturally constituted settings such as ships' bridges, airline flight decks, and science laboratories habitually integrate representations in multiple sense modalities. It must be assumed that these multimodal representations are involved in the construction of memories for past events, the experience of the present, and the anticipation of the future.
- Complex integrated multimodal representations are likely to be more stable than single-mode representations. One way to accomplish this is to embed the representations in durable material media; what I have elsewhere called, "material anchors for conceptual blends." Another way to do this is to embody the representations in bodily motions. These become "somatic anchors for conceptual blends." Stabilization of complex conceptual representations by either means facilitates their manipulation.
- Integrating representations in a single embodied process facilitates sharing the reasoning process with others and the establishment and maintenance

of intersubjective understandings (the so-called common ground).

- Culturally-embedded embodied thinking and acting benefit from adaptive possibilities via both the variability in interactions with material representations and the variability inherent in social interaction. We know least about this aspect of these systems.

Imagining the pre-history of the cognitive life of things
The square-cut gem diagram describes a cognitive ecology. By cognitive ecology I mean that all of the elements and relations potentially interact with one another and that each is part of the environment for all of the others. At this point we must also keep in mind that this diagram is dramatically oversimplified. Most of the elements are hidden in the summary terms. This rich cognitive ecology gives rise to some powerful cognitive processes. The embodied interaction with things creates mechanisms for reasoning, imagination, "Aha!" insight, and abstraction. Cultural things provide the mediational means to domesticate the embodied imagination.

This ecology has been in place – in different forms - for a very long time. The sorts of accounts we create about the development of cognition should be informed by and sensitive to this rich system of interactions. If we were to imagine projecting this interaction system back in time, the elements of the system and the relations among them would surely be changed. Some elements and relations would be changed more than others. If projected back to an ancestor at the level of development of contemporary chimpanzees, for example, the entire square-cut gem of interaction would be present. One big change would be that talk is replaced by other sorts of vocalizations. It is interesting that the super-segmental aspects of vocalization are already present and playing similar roles then as they play now in the expression of emotion. Complex body/world relations are present in tool use and exploration (imagination?) of ways to use tools. I defer to other with more expertise, but I believe that some forms of environmentally-coupled gesture are also present in activities such as threatening with a stick. On the B/B/W facet Chris Johnson has documented learning by socially-prompted exploration of objects such as a stick or a rope (Johnson, n.d.). Body/Body relations are present of course, for this is where animals enact various forms of coordination (greeting, fighting), cooperation (grooming, carrying, mating), and collaboration (group hunting). Body/Vocalization relations exhibit many forms of coordination, although some maintain that chimps do not point (Tomasello, in press). As far as I know, Vocal/Vocal/World relations would be weak. One might observe hooting at the approach of a predator, but reference would presumably be limited.

Now, to move this system forward through time, we should ask, "What happened in the square-cut gem system of interaction during the development of interaction practices?" The possible mechanisms of change are multiplied by the complexity of the relations among the components of the system. Possible explanations of the development of more powerful systems of reference, for example, could be constructed by imagining new systems of embodied coordination such as those documented in the first sections of the paper. Making this move would free the field from the constraint of locating mechanisms of change inside individuals.

Speaking of what is inside individuals, how should one imagine the role of the brain? Before projecting the brain back in time, let's consider its role in the interaction system. In a seminal paper, Rumelhart et al. (1986) argued that people are naturally good at just a few things: matching patterns, manipulating objects in the world, and imagining simple dynamics. We can take these to be properties of the operation of the brain. It is easy to see that these are likely properties of ancient as well as modern brains. Rumelhart et al. go on to say that these abilities account for many of the accomplishments of human cognition, including symbolic behavior. By interacting with particular kinds of cultural things, we can produce complex cognitive accomplishments while employing simple cognitive processes. Once we have learned to interact with these things, we may learn to imagine both the things and our interaction with them. Then we can organize our thinking using internal resources in ways that previously required interaction with external cultural things. In this perspective, interaction with the material and social world come first, and imagination of those interactions come later. This is a thoroughly Vygotskian point of view. But can we give it an embodied interpretation? The bodily interactions with things give rise to somatic anchors for conceptual blends. Imagining bodily interaction with things can become a form of conceptual thinking. As imagined actions become more stylized (remember the gradient of tangibility of representations of LOPs) the somatic anchors for conceptual blends become increasingly detached from the conditions of bodily-engagement-with-the-world that gave rise to them. This could be described as the imagination of embodied abstraction.

Put the Rumelhart et al. perspective together with the square-cut gem model of interaction and project it into the past. As we bring it forward in time, we might now ask, "What changes in brains are required to account for the development of modern interaction practices?" Since there are so many possible mechanisms of change in this complex system, I suspect that the required changes in the brain are less dramatic than has been assumed commonly.

Of course, not all thinking and imagination are carried out in the interaction of body, talk, and cultural world. However, it is a reasonable hypothesis that all thinking and imagination have their origins in this interaction system. It

is for this reason that imagining the cognitive life of things will be essential to understanding cognitive life in general.

Acknowledgments

References

Alač, M. 2003 "Squashing, Rotating, Seeing, and Going: On Visual Knowledge in fMRI Research", *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*, Richard Alterman and David Kirsh (eds.), 2003, 73-78.

Alač, M & E. Hutchins 2004 "I See What You are Saying: Action as Cognition in fMRI Brain Mapping Practice", *Journal of Cognition and Culture*, 4:3, 2004, 629-661.

Becvar, L.A., Hollan, J., Hutchins, E. (in press a) "Representational Gestures as Cognitive Artifacts for Developing Theory in a Scientific Laboratory." *Artifacts in Workplace Practice*. Kluwer Academic Publishers.

Becvar, L.A., Hollan, J., Hutchins, E. (in press b) "Hands as Molecules: Representational Gestures as Cognitive Artifacts for Developing Theory in a Scientific Laboratory." *Semiotica*.

Enfield, N. J.. (2005). The body as cognitive artifact in kinship representations. *Current anthropology*, 46(1), 51-81.

Goodwin, C. 1994 Professional Vision. *American Anthropologist* 96(3): 606-33.

Goodwin, C. (n.d.) Presentation given at UCSD.

Hindmarsh, J.and Heath, C. Embodied reference: A study of deixis in workplace interaction. *Journal of Pragmatics* 32 (2000) 1855-1878.

Hutchins, E. 1995 *Cognition in the Wild*. Cambridge, MA: MIT Press

Hutchins, E. 2005 Material anchors for conceptual blends. *Journal of Pragmatics* 37 (2005) 1555–1577

Hutchins, E. (in press) The distributed cognition perspective on human interaction, In N. Enfield & S. Levinson (Eds.), *Roots of Human Sociality*. Wenner-Gren. Elsevier.

Hutchins, E. & Palen, L. (1997) Constructing Meaning from Space, Gesture, and Speech. In L. B. Resnick, R. Saljo, C. Pontecorvo, and B. Burge (Eds) *Discourse, tools, and reasoning: Essays on situated cognition*. Heidelberg, Germany: Springer-Verlag. Pp. 23-40.

Jefferson, G. 1984. Transcription Notation, in J. Atkinson and J. Heritage (eds), *Structures of Social Interaction*, New York: Cambridge University Press, 1984.

Johnson, C. n.d., Presentation to the UCSD DCOG-HCI lab, March, 2006.

Kirsh, David, 1995. The intelligent use of space. *Artificial Intelligence* 73 (1–2), 31–68.

Murphy, Keith M. 2004. "Imagination as Joint Activity: The Case of Architectural Interaction", *Mind, Culture, & Activity*, 11 (4), 270-281.

Nakahara, Hiroyuki, 1996. Japanese Hand Calculator. Unpublished manuscript, University of Tokyo.

Rumelhart, D., Smolensky, P., McClelland, J., and Hinton, G. 1986. Schemata and sequential thought processes in PDP models. IN J. McClelland, D. Rumelhart, and the PDP research group (eds.), *Parallel Distributed Processing: Explorations in the microstructure of cognition, volume 2*. MIT Press.

Singer, M. A., & Goldin-Meadow, S.A. Children learn when their teachers' gestures and speech differ. *Psychological Science*, 2005, 16, 85-89.

Tomasello, M. (in press). Why don't apes point? In N. Enfield & S. Levinson (Eds.), *Roots of Human Sociality*. Wenner-Gren. Elsevier