# 15 Enaction, Imagination, and Insight

**Edwin Hutchins**

## 15.1 Introduction

Distributed cognition is a framework for exploring the cognitive implications of the commonsense observation that in systems characterized by multiple levels of interacting elements, different properties may emerge at different levels of organization. Thus, a colony of social insects has different properties than any individual insect in the colony (Seeley and Levien 1987; Turner 2000; Holldobler and Wilson 2009). At the level of organisms, bodies have different properties than organs, which have different properties than cells. In the realm of cognition, a neural circuit has different properties than any of the neurons in the circuit. The same can be said of a brain area with respect to the neural circuits that compose it, or of an entire brain with respect to the areas that interact within the brain. This is also true of the body/brain system with respect to either brain or body, and the world/body/brain system with respect to any of its parts. A system composed of a person in interaction with a cognitive artifact has different cognitive properties than those of the person alone (Bruner, Olver, and Greenfield 1966; Cole and Griffin 1980; Norman 1994; Hutchins 1995a, b; Clark 2001, 2008). A group of persons may have cognitive properties that are different from those of any person in the group (Halbwachs 1925; Roberts 1964; Hutchins 1995a; Surowiecki 2004; Sunstein 2006). This layering of scales of integration finds expression in the boundaries among traditional scientific disciplines. More recently developed interdisciplines, of which cognitive science is but one example, search not only for regularities and explanations within levels, but also for patterns in the regularities across levels. The cognitive accomplishments of all human groups depend on the simultaneous operation of cognitive processes on all of these levels from neuron to social group. The big questions in contemporary cognitive science concern the ways that humans, understood as biological creatures, can produce culturally meaningful outcomes.

A central claim of the distributed cognition framework is that the proper unit of analysis for cognition should not be set a priori, but should be responsive to the nature of the phenomena under study. For some sorts of phenomena, the skin or skull of an individual is exactly the correct boundary. For some phenomena, the whole person is just too big and including the whole organism would involve too many interactions. For other phenomena, setting the boundary of the unit of analysis at the skin will cut lines of interaction in ways that leave key aspects of the phenomena unexplained or unexplainable. Most work in distributed cognition to date has focused on systems that are larger than an individual (Hutchins 1995a, b, 2000, 2005, 2006). In these systems, high-level cognitive functions such as memory, planning, decision making, reasoning, error detection and correction, computation, learning, and so on can be identified and analyzed in the culturally organized activities of groups of people in interaction with one another and with technology. Moving the boundaries of the unit of analysis out beyond the skin of the individual human is one important strategy for the distributed cognition approach. It allows us to see how it can be that many of the cognitive accomplishments that have routinely been attributed to individual brains are in fact the accomplishments of cognitive systems that transcend the boundaries of individual bodies. This strategy worked well because the language that classical cognitive science had used to describe internal cognitive processes turned out to be perfectly suited to describing external cognitive processes. Of course, this was no accident. The language of classical cognitive science arose from a distillation of folk observations about external cognitive processes and was given metaphorical extension to the unobservable internal processes (Gentner and Grudin 1985; Hutchins 1995a, chap. 9).

Distributed cognition as applied to socio-cultural systems suggested an answer to the question of how low-level processes create high-level cognition. The idea is that high-level cognition is produced by the culturally orchestrated application of low-level cognitive processes to cultural materials, that is, elements of language, sign systems, and inscriptions of all sorts (Vygotsky 1986; Norman 1994; Hutchins 1995a; Clark 2001).

A simple example of this idea taken from the world of ship navigation is provided by the so-called three-minute rule, which navigators use to compute ship's speed from elapsed time and distance traveled. This instance of high-level cognition computes the value of an abstraction, speed, which is a relationship between distance and time that can be sensed, but cannot be measured directly or expressed with precision by the organic human body. The three-minute rule depends on a serendipitous interaction between two systems of distance units and a system of time units. A nautical mile is very nearly 2000 yards, and an hour is exactly 60 minutes. This means that three minutes is one-twentieth of an hour and 100 yards is one twentieth part of a nautical mile. Thus, the number of hundreds of yards traveled by an object in three minutes equals the speed of the object in nautical miles per hour.[1] This convenient fact is put into practice in navigation in the following way. Two successive positions of a ship are plotted on a three-minute interval. Suppose the distance between them is 1500 yards. The navigator computes ship's speed to be 15 knots by doing the following: "The distance between the fix positions on the chart is spanned with the dividers and transferred to the yard scale. There, with one tip of the divider on 0, the other falls on the scale at a tick mark labeled 1500. The representation in which the answer is obvious is simply one in which the navigator looks at the yard-scale label and ignores the two trailing zeros" (Hutchins 1995a, 151–152). In this analysis, high-level cognitive functions were seen to be realized in the transformation and propagation of representational states. The span between the fix positions on the chart is a representational state that is transformed into a span on the dividers. This representational state is then transformed into a span on the yard scale. Finally, the span on the yard scale is transformed into the answer by reading the label on the designated tick mark in a particular way. Notice that, even though they are obviously involved, in this account, little is said about the use of the eyes, and nothing at all is said about the use of the hands or other parts of the body. In the next section, I will try to show what can be gained by examining the role of the body more closely.

## 15.2 Embodied and Enacted Cognition

Over the past two decades, cognitive science has been shifting from a concept of cognition as a logical process to one of cognition as a biological phenomenon. As more is learned about the biology of human cognition, the language of classical cognitive science, which described external cognition so well, appears increasingly irrelevant to internal cognitive processes. As Clark puts it,

> Perception itself is often tangled up with the possibilities for action and is continuously influenced by cognitive, contextual, and motor factors. It need not yield a rich, detailed, and action-neutral inner model awaiting the services of "central

cognition" so as to deduce appropriate actions. In fact, these old distinctions (between perception, cognition, and action) may sometimes obscure, rather than illuminate, the true flow of events. In a certain sense, the brain is revealed not as (primarily) an engine of reason or quiet deliberation, but as an organ of *environmentally situated control*. (Clark 2001, 95; emphasis in the original)

Embodiment and enaction are names for two approaches that strive for a new understanding of the nature of human cognition by taking seriously the fact that humans are biological creatures. Neither approach is yet well defined, but both provide some useful analytic tools for understanding real-world cognition.

Embodiment is the premise that the particular bodies we have influence how we think. The rapidly growing literature in embodiment is summarized in Wilson 2002, Gibbs 2006, and Spivey 2007. I lack the space needed to sort out the many strands of this literature. Let us simply note here that according to the embodied perspective, cognition is situated in the interaction of body and world, dynamic bodily processes such as motor activity can be part of reasoning processes, and offline cognition is body-based too. Finally, embodiment assumes that cognition evolved for action, and because of this, perception and action are not separate systems, but are inextricably linked to each other and to cognition. This last idea is a near relative to the core idea of enaction.

Enaction is the idea that organisms create their own experience through their actions. Organisms are not passive receivers of input from the environment, but are actors in the environment such that what they experience is shaped by how they act. Many important ideas follow from this premise. Maturana and Varela (1987) introduced the notion of "structural coupling" between an organism and its environment. This describes the relations between action and experience as they are shaped by the biological endowment of the creature. Applying the enaction concept to perception, Noë (2004) says that perception is something we *do*, not something that happens to us. Thus in considering the way that perception is tangled up with the possibilities of action, O'Regan and Noë (2001) introduced the idea of sensorimotor contingencies. In the activity of probing the world, we learn the structure of relationships between action and perception (thus the title of Noë's recent book, *Action in Perception* (Noë 2004). These relationships capture the ways that sensory experience is contingent upon actions. Each sensory mode has a different and characteristic field of sensorimotor contingencies.

One of the key insights of the embodied cognition framework is that bodily action does not simply express previously formed mental concepts;

bodily practices including gesture are part of the activity in which concepts are formed (McNeill 2005; Alač and Hutchins 2004; Gibbs 2006, chap. 4). That is, concepts are created and manipulated in culturally organized practices of moving and experiencing the body. For example, Natasha Myers (2008) described biochemists reasoning about molecular structure by using their bodies to imagine stresses among the parts of a complex molecule. James Watson (1968) reported that he and Francis Crick spent hours cutting out stiff cardboard models of nucleotide pairs and then discovered the double helix of DNA by fitting the pieces of cardboard together. This discovery, like so many (perhaps most) others in science was enacted in the bodily practices of scientists. Similarly, gesture can no longer be seen simply as an externalization of already formed internal structures. Ethnographic and experimental studies (Núñez and Sweetser 2006; Goldin-Meadow 2006) of gesture are converging on a view of gesture as the enactment of concepts (Núñez and Sweetser 2006; Goldin-Meadow 2006). This is true even for very abstract concepts. For example, studies of mathematicians conceptualizing abstract concepts such as infinity show that these too are created by bodily practices. (Núñez 2005; Lakoff and Núñez 2000).

Let us now reconsider the three-minute rule with these general principles in mind. This will show that an embodied analysis of the three-minute rule creates explanatory possibilities that simply have no place in the disembodied analysis presented earlier.

The navigator's first step is to see and apply the dividers to the span of space between the position fixes (figure 15.1). This is a visual activity, but also a motor activity. Techniques for the manual manipulation of the dividers require precise hand-eye coordination. As a consequence of decades of experience, skilled navigators acquire finely tuned habits of action and perception. These include sticking the point of one arm of the divider into the previous fix triangle on the chart, adjusting the spread of the dividers while keeping the point planted, and locating the next fix triangle first visually, and then with the other arm of the dividers. What makes one fix triangle the "previous fix" and the other one the "next fix"? Or, even more basically, what makes a particular set of lines on the chart a fix triangle? The answer to these questions brings us to some fundamental issues concerning interactions with cultural worlds. Many people seem to assume that the status of external representations qua representations is unproblematic. But what makes a material pattern into a representation, and further, what makes it into the particular representation it is? The answer in both cases is enactment. To apprehend a material pattern as a

**Figure 15.1**
Using the dividers to span the distance between successive position fixes.

representation of something is to engage in specific culturally shaped perceptual processes.[2] Regardless of whether the pattern is a sound (apprehended as a word) or a pattern of lines on a chart (apprehended as a position fix), this most powerful of cognitive processes cannot be accomplished any other way.

This fact is expressed differently in different approaches. Goodwin (1994) describes a process by which discursive practices (plotting lines of position, for example) are applied to a domain of scrutiny (a region on a navigation chart) to produce phenomenal objects of interest (a position fix, for example). The label "discursive practices" suggests a narrow a class of perceptual processes that can be so applied. I prefer to say that the enactment of cultural practices in interaction with culturally organized worlds produces the phenomenal objects of interest. In the tradition of phenomenology, the experienced set of phenomenal objects of interest would be referred to as an "own world" (*monde propre*). It is important to notice here that the own world does not consist of isolated objects, but of a system of enacted understandings. The fix is seen as a representation of the position of the ship only when the chart is seen as a representation of the space in which the ship is located. The cultural practices that enact


**Figure 15.2**
Transferring the spanned distance to the scale where the span may be read as either a distance or a speed depending on the way the spanned space is embedded in the navigator's activity.

these understandings may become over-learned and operate outside the consciousness of the person engaging in them.

The navigator's activity at any given moment is embedded in the knowledge of many other moments. The visual appearance of the current span may be compared to other spans that have been plotted. The manual feel of the current span may be compared to other spans or to the largest or smallest distance that can be comfortably spanned with this set of dividers. Once the distance traveled has been spanned with the dividers, a different set of manual skills is required to move the span to the scale (figure 15.2). The navigator must now raise the dividers and move them without changing the span. He must then stick one arm into zero point of the scale, bringing the other arm down to the scale without changing the span.[3]

The activity at any given moment is not only shaped by the memory of past activities, but is also shaped by the anticipation of what is to come.

The navigator's grip on the dividers and the position of his body while spanning the distance on the chart are configured in ways that anticipate moving the span to the yard scale. Thus, experience is not only multimodal, but is also multitemporal or temporally extended in the sense that it is shaped both by memories of the past (on a variety of time scales ranging from milliseconds to years) and by anticipation of the future (over a similar set of time scales).

The activity of using the chart and plotting tools with the three-minute rule involves multimodal experiences in which visual and motor processes must be precisely coordinated. That fact is obvious, but is it relevant? Isn't it safe to disregard these movements of eye and hand as mere implementation details? I believe that we do so at our peril. These embodied multimodal experiences are entry points for other kinds of knowledge about the navigation situation. Bodily experience in the form of unusual muscular tension, for example, can be a proxy for important concepts such as the realization that an atypical distance is being spanned. This implies that sensorimotor contingencies are also learned when the perception of the world is mediated by tools. Chart distances apprehended via the hands and dividers are characterized by a different set of contingencies than distances apprehended visually.

Havelange, Lenay, and Stewart (2003) make an important claim about the difference between human enacted experience and the experience of other animals. In humans, the apparatus by which structural coupling is achieved may include various kinds of technologies.

"We have seen that the own-world of animals is constitutively shaped by the particularities of their means of structural coupling. It is the same for human beings with the enormous difference that the means of structural coupling of humans includes their technical inventions" (Havelange, Lenay, and Stewart (2003, 126; translation by the author). These technologies range from the basic human cognitive technology of language—words are, after all, conceptual tools—to charts and computers and all of the other cognitive artifacts with which humans think. The relevance of this to our current discussion is that a tool—in this case, the divider—is part of the system that produces the particular set of relations between action and experience that characterize the structural coupling of the navigator to his world.

Recent work in embodied cognition suggests that interactions among modes in multimodal representations may be more complex than previously thought. For example, Smith (2005) shows that the perceived shape of an object is affected by actions taken on that object. Motor processes have also been shown to affect spatial attention (Engel, this volume, chapter 8; Gibbs 2006, 61). Thus, we should expect that embodied, multimodal experiences are integrated such that the content of various modes affect one another. Although the sensorimotor contingencies of perceptual modes are distinct from one another, as long as an activity unfolds as expected, the contents of the modes should be congruent with one another. That is, what the navigator sees should agree with what the navigator feels in his hands as he manipulates the tools. The interactions among the contents of various modes of experience will be an important part of the argument to follow.

Once the divider is placed on the distance scale, the navigator uses the pointer of the divider arm to direct his attention to the region of the scale under the pointer. Through this perceptual practice, the divider pointer is used to highlight (Goodwin 1994) a position on a distance scale. The complex cultural skills of scale reading and interpolation produce a number that expresses the value of the location indicated on the distance scale. The scale is perceived in a particular way by embedding that perception in action. What is then seen on the scale is a complex mix of perception, action, and imagination. The cultural practice of speaking or subvocalizing the number expresses the value of the location indicated on the distance scale, and in coordination with the visual and motor experience of the pointer on the scale forms a stable representation of the distance. The congruence of the contents of the many modes of experience lends stability to the enactment of the measured distance.

Notice that what is seen is not simply what is visible. What is seen is something that is there only by virtue of the activity of seeing being conducted in a particular way. That is, what is seen is what is enacted. Even more fundamentally, seeing a line, a set of crossing marks, and the numbers aligned with the marks as a scale of any sort is itself already an instance of enacted seeing. Ingold's (2000) claim that perception is properly understood as a cultural skill fits well with the enaction perspective. The role of enactment of meaning becomes even more evident in the moment when the "distance" scale is *seen* as a "speed" scale, and the distance spanned by the compass/dividers is read as a speed. It is the same scale and similar practices of interpolation are applied to it. But the practice of reading the span on the scale as a speed rather than as a distance is a different practice; a practice that *sees* something different in the very same visual array. In the opening moments of this activity, the span of the dividers is a distance, but the property of being a distance is created by nothing other than the cultural practices of the navigator. As the navigator moves the span toward the yard scale, the span becomes a speed, but again only because that is how the navigator enacts it in that moment. If perception were a passive process, then this same visual array should give rise to the same experience in both moments of perception. But the fact is that reading the span of the dividers on the scale as a speed is a different experience from reading the span of the dividers on the scale as a distance. In this way, cultural practices orchestrate the coordination of low-level perceptual and motor processes with cultural materials to produce particular higher-level cognitive processes. Which higher-level process is produced depends on learned cultural practices as much as it does on the properties of the culturally organized material setting. Under just the right conditions, an enculturated person can place an extent of space on a scale and can read the span there as either a distance or a speed.

Among the points I hoped to demonstrate here are the following: humans make material patterns into representations by enacting their meanings. A phenomenal object of interest in navigation—in this case, the speed of the ship—is enacted in the engagement of the culturally organized world through the cultural practices that constitute the navigator's professional competence. Because the role of the number produced by reading the scale in the navigator's "own world" is the speed of the ship, we can call it an enacted representation of ship's speed. When a triangle of lines on a chart is "seen as" a position fix, or when the chart itself is "seen as" a depiction of the space in which the ship is located, we can also refer to these as enacted representations. These enacted representations involve the simultaneous engagement of perception, action, and imagination. Enacted representations are dynamic, integrating memory for the immediate past, experience of the present, and anticipation of the future. They are multimodal, in the sense that they may involve the simultaneous coordination of any or all of the senses and any modes of action. They are saturated with affect. They are, of course, dependent on the particularities of the sensorimotor apparatus of the organism. The contents of enacted representations are complex multimodal wholes (worlds) rather than isolated objects. Objects are seen (grasped) to be what they are by virtue of the ways they may be engaged by the acting subject.

The emerging picture of the brain as an organ of environmentally situated control is both compelling and problematic. Clark summarized the problem as follows: "What in general is the relation between the strategies used to solve basic problems of perception and action and those used to solve more abstract or higher level problems?" (Clark 2001,135)

Combining the basic embodiment premise that low-level action and perception are inextricably linked (Clark 2001; Noë 2004) with the idea from Havelange, Lenay, and Stewart (2003) that technologically mediated interaction is part of the process of forming enacted representations, opens a new space of possibilities for understanding how high-level cognitive processes can arise in enactment. This paper is an admittedly speculative attempt to sketch out a map of that space of possibilities. If the embodiment premise and the enaction framework are correct, then cognitive processes should be visible in the fine details of the engagement of a whole person with a whole culturally organized world. Whether such an analysis is possible, and if it is possible whether it will help us understand human cognition is at present unknown. In the following sections, I will attempt to perform such an analysis and I hope to show that it does indeed contribute something new to our understanding of the relations between low- and high-level cognition.

## 15.3 An "Aha!" Insight Seen through the Lens of Enaction

Until recently, ship navigation was performed on paper charts using manual plotting tools (Hutchins 1995a). The data on which this analysis is based were originally collected in the early 1980s on the bridge of a U.S. 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 an 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 along the edge of the protractor arm 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 confidence in a fix is inversely proportional to the size of the fix triangle.

I happened to be on the bridge of a large ship, video-recording navigation activities, when, while entering a narrow navigation channel, the 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 bearing. Rather, they were then required to compute the true bearing by adding the corrected magnetic ship's heading to the relative bearing of the landmark (bearing of the landmark with respect to ship's heading). The magnetic compass is subject to two kinds of errors: deviation and variation. The local magnetic environment of the compass can induce small errors, called deviation, that are a function of the interaction between the compass, the ship, and the earth's magnetic field. Deviation errors vary with magnetic heading, are empirically determined, and are posted on a card near the magnetic compass. Magnetic variation is the extent to which the direction of the earth's magnetic field diverges from true north in the local area. The correct equation is: true bearing of the landmark equals compass heading plus deviation plus magnetic variation plus the relative bearing of the landmark (TB = C + D + V + RB). The loss of the gyrocompass disrupted the ability of the crew to plot accurate positions for the ship. The crew explored various computational variations of TB = C + V + RB while plotting thirty-eight lines of position. Then they discovered[4] that a key term, deviation (D), 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.

How can the discovery that this term was missing be explained? The discovery appeared as an "Aha!" insight. In some sense, the "Aha!" insight that this analysis seeks to explain happened just when we would expect it to appear. It happened when the increasing size of the fix triangles led the plotter to explore explanations for the decreasing quality of the fixes. However, neither the navigator's obvious frustration nor the fact that he was looking for something that would improve the fixes can explain the insight. The analysis presented here seeks to reveal the nature of the process by which the plotter examined the fixes and how that process led to the insight that the deviation term was missing. Taken in the context of the computations that the crew was doing, this discovery was, like most creative insights, mysterious. There was nothing in the pattern of computational efforts leading up to the discovery that indicated that the navigators were 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 monomodal (or even amodal) system.[5] A careful examination of the way a navigator used his body to engage the tools in the setting, however, helps to demystify the discovery process, and to explain why and how it happened when it did. The insight was achieved in, and emerged out of, the navigator's bodily engagement with the setting through enacted representations.

Here is a very brief account of the course of events. Lines of position had been plotted to each of three landmarks, but the fix triangle that was produced was unacceptably large. That the triangle was unacceptably large is clear in a comment from the plotter to one of his coworkers. He said, "I keep getting these monstrous frigging god-damned triangles and I'm trying to figure out which one is fucking off!" This also illustrates the emotional character of the experience of these triangles for the plotter. Such a large triangle was clear evidence of the presence of an error somewhere in the process that created the fix. The LOPs were then checked, and at least one possible source of error was tested with respect to each one. These checks did not reveal the source of the problem with the position fix. The plotter then used the plotting tools and the chart to explore changes to LOPs that might improve the position fix. It should be noted that reasoning about the relationships among imagined LOPs is a common practice among navigators (Hutchins 2006). Let's examine this exploration in more detail.

Table 15.1 contains two columns. In the left column are descriptions of the observable actions. In the right column are descriptions of the enactment of the phenomenal objects of interest that can be expected to accompany the observed behavior, given the understanding that enactment is dynamic, multimodal, temporally extended, and affectively colored activity that integrates perception, action, and imagination. I recommend that the reader first read down the left column consulting the accompanying figures to get a sense of the course of action undertaken by the plotter. Once the course of action is clear, the reader will be able to judge the aptness of the descriptions of the enactment. I take the descriptions of the observed activities to be unproblematic. They are based on good quality video with multiple audio streams and informed by an extensive body of background ethnographic information (see Hutchins 1995a). Some of the descriptions of enactment are also straightforward. Some follow directly from the observed activity and others can be inferred and justified by the background ethnography. There are, however, some aspects of the enactment that are clearly speculative. I have marked these in the table with the phrase, "Let us *speculate*."

**Table 15.1**

Observed actions and the hypothesized enactment of phenomenal objects of interest

| Observed activity | Enactment of phenomenal objects |
|---|---|
| The plotter aligned the hoey arm approximately for one landmark, and placed his right index finger on the location of the landmark forming a pivot. He then moved the base of the hoey left, rotating the arm slightly clockwise with respect to the previously plotted LOP for that landmark. This rotation brought the provisional LOP into the interior of the previously plotted triangle, thus reducing the size of the triangle formed with the other two LOPs. | This manipulation of the hoey on the surface of the chart integrates motor, visual, proprioceptive, and tactile experience in an enacted representation of a new LOP. Performed in the culturally meaning space of the chart, this enacts complex conceptual content. Not just a tentative new LOP, but a clockwise rotation, a shift of the LOP to the west-southwest, a smaller triangle, and an improved fix. Examining the placement of the tool on the chart adds stable visual elements to the enacted representation. And these are only the aspects that are demonstrably relevant to the current activity. The navigator must have also experienced the friction of the hoey on the chart surface, the mass distribution of the hoey, and the transparency of the plastic in the hoey arm.<br>These are present in the sensorimotor contingencies of tool manipulation.<br>The tentative nature of this act marks this exploratory manipulation as an example of the class of 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 experience of potential, but not yet realized states of affairs or processes. The fact that these activities are enacted in the subjunctive mood, marked as projecting or anticipating a possible future, is very important. Let us *speculate* that this projection keeps the enacted, embodied anticipation of clockwise rotation active during the following seconds of activity. |
| Observed activity | Enactment of phenomenal objects |
| He then quickly shifted the hoey on the surface of the chart and aligned it approximately with the second landmark, placing his left thumb on the hoey arm near the landmark to serve as a pivot. He also adjusted this LOP slightly clockwise by pulling his right hand and the hoey arm slightly toward his body (figure 15.3). | In these first two moves, the plotter used his body and the tools (chart and hoey) to imagine LOPs that, if they could somehow be created in the future, would make the fix triangle smaller. |

**Table 15.1**
(continued)

| Observed activity | Enactment of phenomenal objects |
|---|---|
| The plotter spoke (self-regulatory speech) the remembered bearing to the third landmark, "one two zero" degrees, while the hoey was still lying on the chart. | Self-regulatory speech enacts the bearing in the verbal modality to form a more stable guide to action. Skilled navigators experience bearing numbers as bodily sensations with respect to a cardinal direction frame. The enactment of the spoken bearing is also embodied in these sensations, and this would have been part of the active context for the next action. |
| Observed activity | Enactment of phenomenal objects |
| The plotter then picked up the hoey in his left hand and used his right thumb to move the arm counterclockwise in the direction of the 120-degree scale position. | Let us *speculate* that the plotter attends visually to the scale values on the protractor in the context of "felt" directions, and the still active enactment of a seen fix triangle and the multimodal anticipation of the small clockwise rotations of LOPs. |
| Observed activity | Enactment of phenomenal objects |
| The plotter made a sharp intake of breath, stopped pushing the hoey arm with his thumb, quickly lowered the hoey held in the left hand to the chart surface, lowered the right hand, which was holding a pencil, to the chart surface beside the hoey, and looked up away from the hoey and chart. All of this happened in less than a second. | This is a clear abandonment of the activity of setting the hoey arm to a scale position, which would have been the first step in plotting the third LOP. Let us *speculate* that the elements of the enacted representations have now combined such that the anticipated multimodal experience of small clockwise rotation is superimposed on the visual experience of the protractor scale (figure 15.4). This combination would produce as an emergent property the concept that adding a small number (small clockwise rotation on the scale) to the bearing for LOP3 will reduce the size of the fix triangle. |

**Table 15.1**
(continued)

| Observed activity | Enactment of phenomenal objects |
| --- | --- |
| The plotter said to himself, “I know what he’s doing!” He tapped the eraser end of his pencil on the chart three times. He then took three actions in quick succession: (1) He turned away from the chart and moved toward the helm station saying, “Let me try . . . Let me try . . . Let me try with my new ones . . .” He consulted the deviation table posted near the magnetic compass at the helm station. (2) The plotter then came back to the chart table, saying, “say three, say three (accompanied these words with beat gestures), add three to everything.” (3) Upon hearing the plotter say this, the bearing timer asked, “Add three? Because we’re shooting relative?”<br>The plotter responded, “Um, no. On a southwest heading add three.”<br>The plotter then re-plotted the three LOPs, adding 3 degrees to each. This produced the desired small fix triangle. | This action sequence contains more self-regulatory speech.<br>Three new concepts have been integrated in this moment. They correspond, in order to the three observed action elements. They are as follows: (1) That the small number that would improve the LOP is deviation. The deviation table is posted at the helm station. (2) That all three LOPs will be improved by adding to them a small number. He says “add three to everything.” (3) That deviation, 3 degrees, is the small number that has been missing from the calculations up to this point. This is clear from the plotter’s statement linking the ship’s heading to the need to add 3 degrees to the LOPs.<br>These three concepts form a synergistic cognitive ecosystem in which each of them makes the others stronger. |

**Figure 15.3**
The positioning of the body of the plotter while adjusting the second LOP slightly clockwise. The left thumb acts as a pivot while the right hand slides the hoey arm slightly toward the plotter's body.

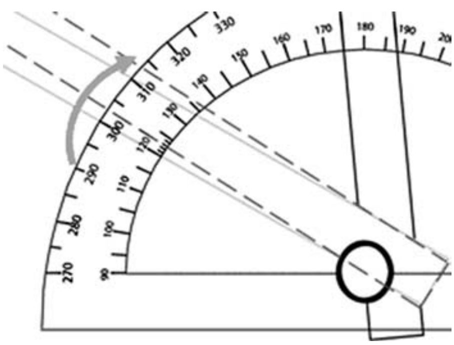


**Figure 15.4**
The superimposition of imagined clockwise rotation (motor anticipation) onto the visual experience of the hoey degree scale. Light-gray solid lines represent the position of the hoey arm when aligned with the 120-degree mark. Dashed lines represent the imagined location of the hoey arm if it were rotated slightly clockwise. The image of a number slightly larger than 120 is an emergent property of this interaction between contents of visual experience and motor anticipation.

There are two speculations here, both of which concern the process of sensorimotor integration. The first is that the enactments of the LOPs produced by the plotter are temporally extended such that anticipatory elements formed early in the process can affect elements that are formed later in the process. The second speculation is that the representations enacted by the plotter are multimodal and that the contents of the various modes may interact with one another. There is ample evidence for the presence of processes that support both of these speculations. First, prediction and anticipation are core functions of animal perception/action systems (Churchland, Ramachandran, and Sejnowski 1994; Noë 2004) and the temporal dynamics of many sorts of action are characterized by both feedforward and feedback effects (Spivey 2007). In fact, the perception of a match between anticipated and current experience even appears to play an important role in an organism's sense that activity belongs to the self (Gibbs 2006). It is therefore plausible that anticipated elements of an enacted representation could interact with elements of subsequent enactments. Second, not only do the contents of various perceptual modes interact with one another, but these interactions have also been linked to success in insight tasks. Spivey (2007, 266–268) describes Glucksberg's (1964) replication of Duncker's (1945) famous candle problem. The problem is to mount a candle on a wall using only the candle, a book of matches, and a cardboard box full of thumb tacks. (The solution is to use the tacks to affix the box to the wall, and use the box as a shelf for the candle.) Glucksberg recorded what the participants did with the actual objects as they attempted to solve the problem. Those who successfully solved the problem tended to touch the box more than those who did not. For those that did solve it, Spivey observes, "Moreover, right before that 'Aha!' moment, the object that these participants had most recently touched was always the box—*and in most cases that touch had been adventitious and nonpurposeful*. It is almost as if the participant's hands suspected that the box would be useful, in and of itself, before the participant himself knew!" (Spivey 2007, 268; emphasis in the original.)

This suggests that the embodied processes of interacting with the material objects may have included the imagination of manipulations of the box that could be useful in solving the problem. More recently, Goldin-Meadow (2006) has shown that children explaining their incorrect answers to arithmetic problems sometimes produce gestures that do not entirely match the contents of their spoken words. In particular, the "gesture-speech mismatches" sometimes highlight with gesture aspects of the correct solution that the student is not yet capable of describing in words. This condition is shown to be an indicator of a readiness to learn the correct solution procedure. Again, reasoning processes playing out in the actions of the hands may hold content that can lead to insights.

The fact that low-level processes can acquire conceptual content when they are deployed in interaction with cultural technology (Hutchins 2005; Havelange, Lenay, and Stewart 2003) suggests that the mechanisms that govern the integration of sensorimotor representations could also shape the integration of conceptual representations. A truly difficult set of questions remain. What principles govern the integration of enacted representations? Do the processes that control the integration of perceptual content also control the integration of conceptual content? Why does cross-modal or cross-temporal integration not destroy representations? These difficult questions need empirical investigation. Ultimately, the answers to these questions will determine the plausibility of the speculations set forth in this paper.

In the fix plotting example, the "Aha!" insight is that the deviation term is missing. The enactment approach gives us a way to see how this insight could emerge from the embodied, multimodal, temporally extended enactment of provisional LOPs that will reduce the size of the fix triangles. The descriptions of the enacted representations I offer earlier are simply what would be expected given the observable behavior of the plotter. No speculation is required to produce the elements from which the solution emerges. The observed enactment of the provisional LOPs includes the experience and anticipation of clockwise rotation of the LOPs. The visual experience of the protractor scale is a necessary component of the activity the navigator is engaged in.[6] The most controversial claim here is that a visual/motor memory of an activity performed in the subjunctive mood a few seconds in the past could somehow combine with current visual/motor perception to produce visual/motor anticipation of activity projected to take place a few seconds in the future. To put that claim in concrete terms: memory for trying out a rotation of the hoey arm on the chart combines with seeing the hoey arm on the scale in a way that anticipates rotating the hoey arm on the scale. I believe that the enactment approach predicts the integration of the particular elements described above in enacted representations. If this does indeed occur, then this instance of "Aha!" insight is no longer mysterious.

In a traditional cognitive explanation of creative insight, one would postulate the entire discovery process in terms of interactions among unobservable internal mental representations. What makes such accounts mysterious is that such internal representations are isolated from the body and world by theoretical fiat. They may be responsive to body/world relations or react to body/world relations, but they are not part of body/world relations. By construing the engagement of the body with culturally meaningful materials in the working environment as a form of thinking, we can directly observe much of the setup for the insightful discovery.

## 15.4 Enaction and Cultural Practices

The processes described thus far 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 2006; Hutchins 2006). In such settings, motion in space acquires conceptual meaning and reasoning can be performed by moving the body. Material patterns can be enacted as representations in the interaction of person and culturally organized settings. Courses of action then become trains of thought. For example, when working on the chart, movement away from the body is conceptually northward, toward the body is south, and clockwise rotation is increasing measure of degrees. When actions are performed by experts in these domains, the integration of bodily sensations with directional frames produces embodied reasoning. Navigators sometimes speak of their reasoning skills in as "thinking like a compass." I believe this could be better described as "enacting compass directions in bodily sensations." The enactments of external representations habitually performed by practitioners who live and work in complex culturally constituted settings are multimodal. It must be assumed that these enacted multimodal representations are involved in the construction of memories for past events, the experience of the present, and the anticipation of the future. Complex enacted multimodal representations are likely to be more stable than single-mode representations (Gibbs 2006, 150). One way to accomplish this multimodal integration is to embed the representations in durable material media—what I have elsewhere called "material anchors for conceptual blends" (Hutchins 2005).

Another way to do this is to enact the representations in bodily processes. These bodily processes become "somatic anchors for conceptual blends." Stabilization of complex conceptual representations by either means facilitates their manipulation. Finally, culturally embedded embodied thinking and acting benefit from adaptive possibilities created by both the variability in interactions with material representations and the variability inherent in social interaction. We know least about this aspect of these systems.

## 15.5 Discussion

From the perspective of a formal representation of the task, the means by which the tools are manipulated by the body appear as mere implementation details. When seen through the lenses of the related stances of embodiment and enactment, these real-world problem-solving activities take on a completely different appearance. The traditional "action-neutral" descriptions of mental representations seem almost comically impoverished alongside the richness of the moment-by-moment engagement of an experienced body with a culturally constituted world. The dramatic difference in the richness of these descriptions matters. Attempts to explain complex cognitive accomplishments using models that incorporate only a tiny subset of the available resources invariably lead to distortions.

The ways that cultural practices adapt to the vicissitudes of situated action are a source of variability in performance, but are often considered to be formally irrelevant to the accomplishment of the task. However, this variability in "task irrelevant" dimensions may be a resource for adaptive processes when routine activity is disrupted.

Multimodality is a fundamental property of lived experience, and the relations among the contents of various modes appears to have cognitive consequences. Goldin-Meadow (2006) proposes a single dimension of variation in the relations between gesture and speech. The contents of these two modes (of course, each, by itself, is richly multimodal) can carry roughly the same information and be matching, or they can carry different information and be mismatched. However, the space of possible relations is larger that this. The contents of gesture and speech can match or mismatch in several ways. Let us call the match condition a case in which the contents of the modes are congruent. The condition that Goldin-Meadow calls "mismatch" could better be described as complementary.

The contents differ, but they differ in ways that can combine to make a single coherent concept. The contents of gesture and speech could also be contradictory, or they could be incongruent in the sense that they are simply irrelevant to each other. Congruence among the contents of modes appears to lend stability to the enacted representations of which they are a part. Complementarity among the contents of modes may give rise to emergent phenomena, as was the case with the "Aha!" insight described in section 15.3 (see also Hutchins and Johnson 2009). Contradictory contents are sometimes produced deliberately in sarcasm. Truly incongruent contents probably occur, but it will be difficult to know how frequently this happens. Incongruent contents will most likely go unnoticed, or, if noticed, will be dismissed as noise.

The enaction perspective reminds us that perception is something we do, not something that happens to us. And this is never truer than when a person perceives some aspect of the physical world to be a symbol or a representation of any kind. Everyone agrees that perceiving patterns as meaningful is a human ability. But as long as perception was conceived as something that happened to us, it was possible to ignore the activity in the world that makes the construction of meaning possible. And although the enaction of cultural meanings is something that our bodies and brains *do* in the world, it is not something that our bodies or brains do by themselves. The skills that enact the apprehension of patterns as representations are learned cultural skills.

Putting things together this way reveals new analytic possibilities for understanding interactions of whole persons with the material and social worlds in which they are embedded. Learned cultural practices of perception and action applied to relevant domains of scrutiny enact the phenomenal objects of interest that define activity systems. High-level cognitive processes result when culturally orchestrated low-level processes are applied to culturally organized worlds of action.

Every mundane act of perception shares something fundamental with creative insight; the fact that what is available to the senses and what is experienced can be quite different. Reading the same scale for distance or speed in the use of the three-minute rule is a simple example. Similarly, a navigator can read the 120-degree mark on the protractor scale as a stable target on which one can position the hoey arm. Or the same navigator might read the same mark as a referent with respect to which a small clockwise rotation produces a new target, a slightly larger number on the scale, that fits better the anticipated course of action. In reading the mark this way, he suddenly sees what had been hidden. "Aha! Add three to everything." What makes ordinary acts of perception ordinary is only that the cultural practices of enacting them are over-learned and the outcomes follow as anticipated. Creative acts of perception can occur when emergent relations arise in the enaction of integrated, multimodal, temporally extended, embodied representations.

## Acknowledgments

The Santa Fe Institute's program on Robustness in Social Systems provided funding for the work reported here. Erica Jen served as grant monitor. Alisa Durán transcribed the data and helped me focus the analysis on the problem of insight. Figure 15.3 was drawn from a video frame by Whitney Friedman. I am grateful to Andy Clark, Kensy Cooperrider, Deborah Forster, Charles Goodwin, Rafael Núñez, and John Stewart for valuable comments on an earlier version of this chapter.

## Notes

1. Virtually all ship navigators know this rule and can use it, but few know why it works.

2. For my purposes, a practice will be labeled cultural if it exists in a cognitive ecology such that it is constrained by or coordinated with the practices of other persons.

3. Notice that the two tasks, adjusting the span, followed by maintaining the span while moving it, put conflicting demands on the tool. It must be mutable one moment, and immutable the next. This problem is solved for dividers by an adjustable friction lock. In fact, friction locks are common, and it is likely that wherever a friction lock is present, embodied knowledge is at work.

4. Other verbs that might be placed here include "noticed" and "remembered." Each implies something about the nature of the process. "Notice" highlights the aspect of happenstance. "Remember" highlights the fact that this is something that all navigators already know. "Discover" emphasizes the fact that they were searching for something that would improve the quality of the fixes when they became aware that D was missing. Including the previously missing D term did improve the fixes and thus ended their search.

5. In Hutchins 1995a, I provide a disembodied analysis of this event that fails to explain how the discovery of the missing term was made.

6. Of course, we cannot conclude anything about the quality of that visual experience from the available data.

## References


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

Bruner, J., Olver, R., and Greenfield, P., eds. (1966). *Studies in cognitive growth: A collaboration at the Center for Cognitive Studies*. New York: John Wiley and Sons.

Churchland, P. S., Ramachandran, V. S., and Sejnowski, T. J. (1994). A critique of pure vision. In *Large-scale neuronal theories of the brain*, ed. C. Koch and J. Davis, 23–60. Cambridge, MA: MIT Press.

Clark, A. (2001). *Mindware: An introduction to the philosophy of cognitive science*. Oxford: Oxford University Press.

Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. Oxford: Oxford University Press.

Cole, M., and Griffin, M. (1980). Cultural amplifiers reconsidered. In *The social foundations of language and thought: Essays in honor of Jerome Bruner*, ed. D. Olson, 343–364. New York: Norton.

Duncker, K. (1945). On problem solving. *Psychological Monographs* 58 (5): 1–270.

Enfield, N. (2006). Social consequences of common ground. In *Roots of human sociality: Culture, cognition and interaction*, ed. N. J. Enfield and S. C. Levinson, 399–430. Oxford: Berg Publishers.

Gentner, D., and Grudin, J. (1985). The evolution of mental metaphors in psychology: A 90-year retrospective. *American Psychologist* 40 (2): 181–192.

Gibbs, R. (2006). *Embodiment in cognitive science*. New York: Cambridge University Press.

Glucksberg, S. (1964). Functional fixedness: Problem solution as a function of observing responses. *Psychonomic Science* 1:117–118.

Goldin-Meadow, S. (2006). Meeting other minds through gesture: How children use their hands to reinvent language and distribute cognition. In *Roots of human sociality: Culture, cognition and interaction*, ed. N. J. Enfield and S. C. Levinson, 353–373. Oxford: Berg Publishers.

Goodwin, C. (1994). Professional vision. *American Anthropologist* 96 (3): 606–633.

Halbwachs, M. (1925). *Les cadres sociaux de la memoire*. Paris: Albin Michel.

Havelange, V., Lenay, C., and Stewart, J. (2003). Les représentations: mémoire externe et objets techniques. *Intellectica* 35:115–131.

Holldobler, B., and Wilson, E. O. (2009). *The superorganism: The beauty, elegance, and strangeness of insect societies*. New York: Norton.

Hutchins, E. (1995a). *Cognition in the wild*. Cambridge, MA: MIT Press.

Hutchins, E. (1995b). How a cockpit remembers its speeds. *Cognitive Science* 19:265–288.

Hutchins, E. (2000). The cognitive consequences of patterns of information flow. *Intellectica* 1 (30): 53–74.

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

Hutchins, E. (2006). The distributed cognition perspective on human interaction. In *Roots of human sociality: Culture, cognition and interaction*, ed. N. J. Enfield and S. C. Levinson, 375–398. Oxford: Berg Publishers.

Hutchins, E., and Johnson, C. (2009). Modeling the emergence of language as an embodied collective cognitive activity. *Topics in Cognitive Science* 1: 523–546.

Ingold, T. (2000). *The perception of the environment: Essays in livelihood, dwelling, and skill*. Oxon, UK: Routledge.

Lakoff, G., and Núñez, R. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York: Basic Books.

Maturana, H., and Varela, F. (1987). *The tree of knowledge: The biological roots of human understanding*. Boston: Shambhala.

McNeill, D. (2005). *Gesture and thought*. Chicago: University of Chicago Press.

Murphy, K. (2004). Imagination as joint activity: The case of architectural interaction. *Mind, Culture, and Activity* 11 (4): 267–278.

Myers, N. (2008). Molecular embodiments and the body-work of modeling in protein crystallography. *Social Studies of Science* 38 (2): 163–199.

Noë. A. (2004). *Action in perception*. Cambridge, MA: MIT Press.

Norman, D. (1994). *Things that make us smart: Defending human attributes in the age of the machine*. Boston: Addison-Wesley.

Núñez, R. (2005). Creating mathematical infinities: The beauty of transfinite cardinals. *Journal of Pragmatics* 37:1717–1741.

Núñez, R., and Sweetser, E. (2006). With the future behind them: Convergent evidence from Aymara language and gesture in the crosslinguistic comparison of spatial construals of time. *Cognitive Science* 30:401–450.

O'Regan, J. K., and Noë, A. (2001). A sensorimotor approach to vision and visual consciousness. *Behavioral and Brain Sciences* 224 (5): 939–973.

L

Roberts, J. (1964). The self-management of cultures. In *Explorations in cultural anthropology: Essays in honor of George Peter Murdock*, ed. W. Goodenough, 433–454. New York: McGraw-Hill.

Seeley, T., and Levien, R. (1987). A colony of mind: the beehive as thinking machine. *Sciences* 27 (4): 38–43.

Smith, L. (2005). Action alters shape categories. *Cognitive Science* 29 (4): 665–679.

Spivey, M. (2007). *The continuity of mind*. Oxford: Oxford University Press.

Sunstein, C. (2006). *Infotopia: How many minds produce knowledge*. Oxford: Oxford University Press.

Surowiecki, J. (2004). *The wisdom of crowds*. New York: Doubleday.

Turner, J. S. (2000). *The extended organism: The physiology of animal-built structures*. Cambridge, MA: Harvard University Press.

Vygotsky, L. S. (1986). *Thought and language*. Cambridge, MA: MIT Press.

Watson, J. (1968). *The double helix: A personal account of the discovery of the structure of DNA*. New York: Simon and Schuster.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review* 9 (4): 625–636.