# A New Cognitive Ethnography

Hutchins, Edwin [1]

[1] *Department of Cognitive Science, University of California San Diego*

ABSTRACT

*A paradigm shift is underway in the behavioral sciences.*

*Recent advances in computational theories of mind are reshaping the theoretical foundations of cognitive science. Deep learning and predictive processing are driving incredible advances in the capabilities of artificial intelligence as well as providing a theoretical basis for the unification of many previously dispersed sub-fields.*

*Over the past five decades, I have developed a method for the study of cognition in the wild. This method identifies empirical observations of the fine-scale details of naturally occurring ongoing activity as instances of theoretical concepts. As theories of cognition evolved and as the sites of my ethnographic fieldwork changed, the cognitive ethnography method continued to bring cognitive theory to the interpretation of real-world activity.*

*Theorizing and speculation about how real-world cognitive processes might appear if the mind were assumed to be a predictive processing system have already begun. Where previous work has theorized the applicability of predictive processing to real-world cognition, the analyses presented here offer a first empirical glimpse of what cognition viewed through this framework actually entails in the fine structure of naturally occurring activity.*

*In this paper, I propose a novel theoretical framework that combines predictive processing with multimodality and the continuous coupling of organism to environment. I use this framework as the underlying theory in the cognitive ethnographic analysis of two dozen vignettes of real-world activity.*

# 1.   Introduction

A paradigm shift is underway in the behavioral sciences.

For decades, the symbolic computational model of the mind dominated cognitive science. It had a theoretical foundation called the Physical Symbol System Hypothesis. This view had a behavioral aspect known as information processing psychology and a computational aspect called symbolic artificial intelligence.

A model of mind based on what has been known variously as parallel distributed processing, connectionism, or neural networks was originally conceived around the same time as the symbolic computational model of mind, but was considered by many to be unworkable (Minsky & Papert, 1988). For decades, connectionism developed in parallel with symbolic AI. Over the years, connectionist approaches went by a variety of names: parallel distributed processing (PDP) (Rumelhart & McClelland, 1986), neural networks, and now Deep Learning and Generative AI.

The latest version of the connectionist approach has a behavioral aspect now known as predictive processing (PP), and a computational aspect known as Generative Artificial Intelligence (genAI). In recent years, genAI has outperformed symbolic AI on a wide range of tasks that are taken to be signs of human intelligence. Some researchers believe that the behavioral aspect, predictive processing, promises to unite neuroscience, which has always occupied the core of cognitive science, with a number of approaches that have previously resided on the periphery of the field (Clark, 2016). The computational aspect, genAI, promises or threatens to profoundly change our world for good or ill, or both good and ill.

Generative AI has recently become the future of technology. Will Predictive Processing become the future of behavioral sciences? How might these developments change the way we imagine human cognition?

There is a potential conceptual trap in these developments. Much of cognitive science has subscribed to the view that "the mind is in the brain", or "the mind is what the brain does". These ideas are pervasive in today's discussions about genAI and its relationship to human intelligence. In the public imagination, genAI is a model of the brain, and the individual brain is the seat of human cognition and cognitive accomplishment. We are warned that some genAI systems will soon be more intelligent than any person, and that this presents a perhaps existential threat to human civilization. GenAI may well present an existential threat to humanity, but it is not because of its relationship to individual cognitive function. As is well known, and as I will show again in the pages below, what makes humans smart is not simply the properties of individual brains. Because neural nets are inspired by the organization of neurons in a brain, the temptation is strong to repeat the mistakes of the past and reduce intelligence to the function of a brain. However, this moment in time presents us with an opportunity to avoid the trap of the conceptual collapse of mind to brain. Perhaps it seems paradoxical

that a better model of brain function can help us avoid attributing too much to the brain. I hope that by the end of this paper, you will agree that this is the case.

## 1.1. My Goal

I intend this paper as an attempt to seize this opportunity; to show how the predictive processing approach implies that the mind transcends the boundaries of the individual brain. I will describe a new framework that incorporates predictive processing to inform and transform our ways of looking at real-world activity. Because I'm a cognitive ethnographer trying to detect, highlight, illuminate, and perhaps explain the cognitive aspects of everyday human behavior, I need a slightly different framework from the usual predictive processing model of brain function. In addition to the generative aspects of PP, my observations of real-world activity make clear that people's engagement with the world is multimodal and continuous. I combine these features to create what I call the Multimodal, Generative, Continuous (MGC) approach. I propose to explore human activity seen through the lens of this generative model of perception, action, and thought. This MGC model is made possible in part by the rise of PP as a demonstration of a plausible model of the generative aspect of individual cognitive processing, but it goes beyond other approaches by insisting on multimodality and the continuity of coupling between an organism and its environment.

## 1.2. The Role of Theory and Researcher Imagination

When we as researchers confront phenomena, we always do so using a network of assumptions. If we are paying attention, we may ask ourselves, "What do our assumptions permit us to imagine?" If our theoretical assumptions indicate that a phenomenon seems unlikely or not supported by theory, then we are unlikely to pursue it. If I believe I am seeing phenomenon x in real-world activity, but x is implausible in the theoretical framework that I assume to describe human cognitive function, a tension is created. I might ask myself, "Am I wrong about the phenomenon, or do my assumptions not serve my needs?" For decades, I have been struggling to find plausible theoretical descriptions for several phenomena that appear in my observations in the field. My excitement and enthusiasm for the MGC framework stems from the fact that

many observable phenomena that formerly seemed implausible follow naturally from the MGC framework.

The MGC framework permits us to imagine cognition working in ways that fit fine-scale observations of ongoing human (and non-human) activity. I know that I am not alone in struggling to find a fit between observation of everyday life and cognitive theory. Other approaches that have developed in the past two decades to address what happens at the interface of organism and environment experience similar tensions.

These tensions are among the forces driving the paradigm shift in the behavioral sciences that I pointed to in my opening sentence. Predictive processing (and I think even more so, the MGC approach) promises to unite many fields that have, until now, remained on the periphery of mainstream cognitive science. Consider ecological psychology (Gibson, 1986; Kugler & Turvey, 1987), dynamical systems theory (Kelso, 1995; Port & van Gelder, 1995; Spivey, 2007; Thelen & Smith, 1994), embodied cognition (Gibbs, 2006; Johnson, 1987; Lakoff & Nuñez, 2000), embodied robotics (Beer, 2008), enaction (Maturana & Varela, 1987; Stewart et al., 2010; Thompson, 2007), predictive processing (Friston, 2008), extended cognition (Clark & Chalmers, 1998), ecology of mind (Bateson, 1972), distributed cognition (Hutchins, 1995a), situated action (Suchman, 1987), cultural-historical activity theory (Cole, 1996; Vygotsky, 1978; Wertsch, 1985), actor network theory (Latour, 2005), and *le coursd'action* (Theureau, 2015). All these approaches attempt to put the focus of study on the interactions of agents with their social and material environments. These approaches consider sensory and motor processes to be part of the apparatus of thought rather than peripheral devices that deliver the world to a central processor in the form of symbolic representations. None of these approaches is well accounted for by the traditional computational model of mind. They have chaffed under the weight of an old model that makes it difficult to imagine the phenomena they study. This includes phenomena that exist in a system that comprises brain, body, and world, and that cannot be understood without considering all the components in interaction.

In this paper, I want to sketch a new edifice of assumptions to guide our thinking about cognition in human activity. I will illustrate the utility of these assumptions by using them to analyze vignettes from everyday life. The analyses presented here are exploratory in the sense that I'm trying to see what is highlighted in human activity when we assume that people are, among other things, MGC systems. This has not been done before. We do not know in advance that such analyses will yield valuable results.

## 1.3. Researcher Attention

Theoretical frameworks always shape what we see. PP systems learn the flow of experience in the sensory and motor systems, so that's where our attention is directed. Generative AI models develop internal processes that can produce, predict, and imagine the course of their experience, but when there are $10^{12}$ parameters in a computational network model, speculation about the details of internal processes is intractable and, for our purposes, unnecessary. At present, we can describe the general principles of how the PP system (and supposedly the brain) does its work, but the details of how any particular task is accomplished remain beyond our reach. The opaqueness of the depths of deep learning systems channels our attention toward the interaction of an organism with its environment. In earlier approaches, attention was focused on internal processes, and the stuff outside the brain was left blurry - literally out of focus. I now propose to invert that situation. The observable actions in, and interactions with, the cultural world are brought into sharp focus, and internal processes may be present in peripheral vision but remain out of focus. PP is useful for me partly because of this effect. Incorporating PP in my theoretical base directs attention from unknowable internal processes to the observable behavior of people in real world activity.

## 1.4. Constraints on the Stories We Tell

It can be liberating to feel no obligation to imagine in detail the internal machinery of thought. But if that is no longer our story, what stories shall we tell? And where does one find the constraints to which these stories must be responsible? I will try to show how constraints on models of cognitive function can be derived from observations of behavior contextualized by deep cognitive ethnography.

In some cases, examinations of ongoing activity in terms of MGC will lead to speculations that cannot be fully evaluated on the basis of observations in the field. In those cases, I hope to raise questions that can be investigated by other cognitive scientists using other methods.

## 2. Background: Inventing Cognitive Ethnography

Let me provide some autobiographical background to my project.

I am a cognitive scientist, but I was originally trained as an anthropologist. In an analogy of cognitive science with biology, I would position myself as a naturalist. That is, while most of my colleaguesstudy cognition in laboratory settings, I have always studied cognition as it occurs in the wilds of everyday life.

Every proper study requires a theory, a method, and an object of scrutiny. As a naturalist of human activity, I put these pieces together as follows: The theory directs our attention. It says what to look for and how to look. The method captures events in the object of scrutiny and identifies them as instances of concepts in the theory.

## 2.1. Reasoning in Trobriand Island Discourse

1

When I started out in the 1970s, it was widely believed that primitive technology implies primitive thought. The fact that illiterate third-world adults performed on IQ tests about like 10-year-old children in Europe or North America was a robust finding in cross-cultural psychology (Dasen, 1972). In the summer of 1973, I was one of three graduate student researchers under the direction of UCSD anthropology professor Theodore Schwartz, who went to Manus Island in Papua New Guinea to conduct such a study. We administered a battery of intelligence tests to adult non-literate farmers and fishermen in several villages. We replicated the expected finding. Many people concluded from studies like these that non-literate unschooled adults in the third world think like children. But something seemed wrong to me. In everyday interaction, these people did not seem in any way mentally impaired. Furthermore, when the children from these same villages went away to school, they scored about like their age contemporaries in the Western world. There was clearly nothing wrong with the brains of the people. What was going on?

With an intelligence test of the sort our team used in Manus, the theory concerns mental abilities. The object of scrutiny is the mental abilities of the subject(s) and the test is a method to capture aspects of those abilities and assign them as instances of the concepts in the theory, i.e., inference, problem solving, abstract thinking, etc. When researchers take an intelligence test into the field, they put the burden on the subjects to figure out the nature of the tasks presented by the test. One can never be sure how much of the

observed performance on the test is due to the abilities of the subjects and how much is due to their attempts to understand the nature of the tasks. Since every culture sets for its members many cognitive tasks, I thought perhaps we should, instead, observe real world cognitive activities, and put the burden on the researcher to figure out the nature of the task. We could still make inferences about cognitive capabilities by analyzing performance in the task. That is, with a sufficiently thorough understanding of the task, observed behaviors may be identified as instances of concepts in a theory of cognition.

Funded by the Social Science Research Council, I went back to Papua New Guinea in 1975. Following in the footsteps of the great ethnographer, Bronislaw Malinowski, I went to the Trobriand Islands to do my PhD dissertation research. The Trobriand Islands was an apt choice because some authors, working from Malinowski's transcriptions of magic spells, had concluded that the Trobriand Islanders did not think logically (Lee, 1949). Once installed in a village, I searched for a naturally occurring activity that would allow me to assess reasoning abilities. I settled on land litigation. Because land litigation is conducted in public, I could record performances in the task. Land litigation is clearly cognitive because litigants must construct a narrative history of a piece of land that results with them holding rights, while simultaneously heading off anticipated counter arguments of their opponents. And because, on a small coral island, nothing is more important that rights in land, the motivation to take the task seriously is built into the activity.

**Figure 2.1. The author recording land litigation in a Trobriand Island village in 1976**

*Photo by Dona Hutchins*

My theory was a combination of propositional logic and schema theory, as it was then developed in cognitive science (Rumelhart, 1975). Schema theory posits underlying conceptual structures that bring organization to knowledge and belief. A schema for buying, for example, includes roles for a buyer, a seller, an item to be bought, and a price paid. The schema describes relations among these roles. The buyer pays the price to seller in exchange for the item. The schema also specifies logical relations. The payment of a price is a necessary condition for the transfer of an item. This can be expressed as a premise: The transfer of the item implies that a price was paid. Binding actual events and things in the world to the roles in the schema produces propositions that may be assigned a truth value. "John paid Jane $2" is a proposition, which might be true or false. If the proposition "John paid something to Jane" can be shown to be false, it can be inferred that no item was bought by John from Jane. T implies P. P is observed to be false. Therefore, T is false. This is an example of *modus tollens* inference.

Adapting to the needs of the situation on the ground, I developed a first draft of an approach that I would later call *Cognitive Ethnography*. This is regular ethnography, as practiced by all sorts of anthropologists, plus some kind of recording of naturally

occurring activities, followed by micro-analysis of the fine-scale details of the recorded activity.

Cognitive Ethnography begins with the classical method of anthropological fieldwork and participant observation. To do this sort of ethnography, it is essential to gain competence in the language that people speak in their everyday lives. Trobriand Islanders speak an Austronesian language known in the literature as Kilivila or Boyowan. My wife and I created a dictionary and a grammatical sketch of the language. Our dictionary is available online. Prior to our work, the only available materials were short word lists assembled by missionaries. Since then, a comprehensive grammar and dictionary has been published by Gunter Senft (Senft, 1986). In addition to writing field notes and taking still photographs, I made audio recordings of land litigation as it occurred in the village where I lived. Through extensive interviewing and a failed attempt to make my own garden, I documented gardening practices and the principles of land litigation.

I transcribed audio recordings in the Trobriand Island vernacular. I then performed a careful micro-analysis of the discourse, identifying statements in the data as instances of the concepts in the propositions that constitute schemas for transfer of rights in land. This permitted me to determine the logical relationships among the utterances produced by the litigants in their public discourse, and to identify the inferences using the typology of propositional logic. From this analysis, I was able to conclude that the Trobriand Islanders make the same kinds of inferences that we make. They are just as logical (or not) as your average American or European. This work was published in my book, *Culture and Inference* (Hutchins, 1980).

Propositional logic and schema theory produce re-descriptions of observable behavior. They permit us to identify inferences in the discourse as instances of strong inference (modus ponens and modus tollens) and plausible inference (affirmation of the consequent and denial of the antecedent), but they provide only weak constraints on the nature of the internal states and processes of the litigants. The observed behavior might have been produced by some internal symbol processing apparatus, or by some other unknown process. Using the language of the time in cognitive science, one would say that the observed data were rule-described, but not necessarily rule-generated. And that is fine. *It is not always necessary for the theory to be a true description of internal processes, nor for the data to tightly constrain internal processes, to make useful assertions about cognitive processes.* Inferences made in discourse are available to direct

observation and can be placed in the typology of propositional logic as long as one knows the schemas in use and knows the particulars of the subject of the discourse sufficiently well to instantiate the schemas as propositions with truth values. Of course, documenting the schemas and knowing the particulars of each case requires extensive ethnographic investigation.

While I did not know what sorts of internal processes produced the observed behavior, I did suspect that the processes, whatever they were, must be quite general. At the time, I called the processes the cultural code. After a discussion of the many kinds of work the code seemed to do, I said,

> "Terms such as problem solving, planning, understanding, decision making, and explanation are often taken as descriptions of cognitive processes. In light of the uses of the cultural code observed in the previous chapter, I take them not to be descriptors of various distinct processes so much as descriptors of the conditions under which, or the task environments in response to which, the cultural code (as a process) is applied." (Hutchins, 1980, p. 110)

This speculation, made in the late 1970s, appears prescient in the light of claims made in the past decade concerning the operation of the brains conceived as predictive processing systems.

This project on Trobriand Island land litigation gave rise to the recipe for all my subsequent cognitive ethnographic research: Begin with traditional ethnography. Choose a small naturally occurring activity. Record data on the performance of the activity. Analyze the data using whatever theory seems best to reveal the cognitive aspects.

Recording technology imposes strong limitations on what is possible with such a method. Malinowski could not have done this project because it requires the analysis of the words people actually say, and in 1916 he had no way to record ongoing speech. Taking notes on what people say in ordinary conversation is not enough. In fact, the only sort of discourse that Malinowski could reliably capture verbatim was magic. Magical spells can be recorded accurately in written notes because the effectiveness of a magical spell depends on the fidelity of its repetition. A spell must be recited exactly in

order to achieve the desired results. A magician informant can be asked to repeat a spell as needed to get it down in writing. Malinowski did transcribe, translate, and publish many magical spells that were recited for him by Trobriand magicians (Malinowski, 1965). Unfortunately, magic has a very different logical structure from everyday discourse. Normal cause-and-effect relations are distorted in magic by the fact that magic is, well, magical. This means that attempting to assess Trobriand Islanders' reasoning abilities by analyzing the language of magic will not produce accurate results.

Of course, there is so much more to the performance of public litigation than the logical structure of the discourse. The audio recording also captures prosody, and other properties of the verbal stream, as well as other sounds such as the wind blowing, rain falling, dogs barking and roosters crowing in the village. Audio, however, cannot capture facial expression, gesture, body posture, the sun beating down or covered by cloud, the coming and going of participants, and smoke drifting through the village.

This illustrates how a theoretical framework, together with data collection apparatus, forms a filter and a spotlight that selectively highlight certain aspects of the phenomena while disregarding or completely failing to see others.

# 2.2. Ship Navigation

2

In the early 1980s, I had a position with the US Navy as a personnel research psychologist (the Navy didn't seem to know how to hire an anthropologist).

At first, I rode ships in the engineering spaces doing ethnography on the operation of steam propulsion systems in support of the development of the STEAMER computer-based instructional system (Hollan, Hutchins, & Weitzman, 1984). On subsequent field trips I moved to the Combat Information Center where I worked on radar navigation systems. There I was able to discover and document how an innocent looking deviation from procedures caused major problems for task-force coordination. In those days, operations specialists in the CIC used a polar-coordinate plotting sheet called the maneuvering board to compute important features of the ship's relationship to other ships, such as closest point of approach, collision threat, scouting tracks, and so on. Naval training centers experienced a high rate of failure in courses teaching the use of

the maneuvering board. I designed a computer-based training system for use of the maneuvering board that reduced the failure rate to about one tenth of its previous value. This system was subsequently adopted as standard training aboard every ship in the US Navy (Hutchins & McCandless, 1982).

Neither the engineering spaces nor the Combat Information Center have windows through which one can see the world outside the ship. In all my observations to this point, I had not yet experienced either of the two most romantic moments aboard ship, the departure from and the arrival at a harbor. Finally, at the conclusion of a trip from San Diego up the west coast aboard the amphibious transport dock ship, USS Denver, I decided to go onto the bridge for the entry into the Straits of Juan de Fuca and the arrival in the port of Seattle. The bridge was busy, and the activity of the navigation team was simultaneously familiar and new to me. Having been trained as a navigator of offshore racing yachts and being a member of the last generation of celestial navigators, I knew a fair bit about navigation, but nothing at all about how it was done on ships. After observing the navigation activity on the bridge of the USS Denver, I resolved to make a more focused study of bridge navigation.

My theoretical framework was the symbolic computational model of mind. This theory postulated a central cognitive processor that did the thinking by manipulating strings of symbols. Sensory systems transformed sense data into symbols to be passed to the central processor. The central processor could pass symbols to motor systems that transformed strings of symbols into movement. Newell and Simon called such a system the "Physical Symbol System," and proposed in the Physical Symbol System Hypothesis (PSSH) that wherever intelligence is found, it will be found to be a PSS (Newell & Simon, 1972).

Thus, armed with the dominant theoretical frameworks in cognitive science at the time and a good start on the ethnography of Western navigation, I set out to learn about the cognition of individual navigators on Navy ships. I began making field observations on the bridge of several ships.

**Figure 2.2. The navigation team at the chart table on the bridge of the USS Pala**

*Photo by the author*

I experienced an epiphany one afternoon while standing on the bridge of a ship watching the navigation team work as the ship entered San Diego harbor. It suddenly became clear to me that the outcomes that matter to the ship, such as whether it went aground or not, were determined by the team rather than by any individual navigator. Writing up my field notes later that night, I speculated that the navigation system, comprised of four quartermasters plus their tools, might have cognitive properties of its own.

It came as a surprise to discover that the PSSH provided an excellent metaphorical account of what was happening among the members of the team. They were creating both symbolic and non-symbolic representations of the situation of the ship and propagating and transforming those representations. There was a flow of information through the system. There were long-term memory stores in the form of charts and short-term memory stores in the form of logbooks and pencil marks on the charts. There were analog-to-digital transformations implemented in the telescopes used for observing bearings of landmarks and digital-to-analog transformations implemented in the plotting tools.

By the time I wrote *Cognition in the Wild,* eight years after my epiphany on the bridge, I understood that the fact that the activities on the navigation team fit perfectly with the Physical Symbol System Hypothesis was no accident. Contrary to the belief in cognitive science at the time that the computer was made in the image of the human (Simon & Kaplan, 1989), it was clear that the computer, which implemented a PSS, was made in the image of a socio-technical system like the navigation team. I made this argument in the last chapter of *Cognition in the Wild* (Hutchins, 1995a).

The sort of symbolic cognition implemented by the navigation team and its tools captures something important about humans. It is the secret of our success as a species. We represent the world in external symbolic expressions such as equations. We transform the symbols using rules that respond only to the form of the symbols, not to their meaning. And then we re-interpret the new symbolic expressions as descriptors of states of the world. This process of symbolic representation and manipulation is the foundation of logic and mathematics, science, and engineering. It allows us to predict the future and describe that which cannot be directly observed. It gives us access to and control over the absent and the abstract.

Looking back over 40 years, I now see that armed with the core concepts of classical symbol-processing AI and information processing psychology, I went looking for physical symbol systems on the bridge of a ship. And I found them. I honestly expected to find them inside the navigators, but they were not where they were expected! They were not, as far as I could tell, inside the navigators. They were, instead, between the navigators, or among the navigators and the culturally elaborated task setting in which the navigators worked. What makes us smart? It's more than big brains. Humans create their cognitive powers by creating the physical and social environments in which they exercise those powers.

This epiphany was the origin of the approach I came to call "*distributed cognition.*" Here was an observable system that manifested precisely the features called for by the PSSH. But it was not located in a central symbol processor, it was *distributed* across the members of the team, the devices and artifacts they manipulated, the procedures they followed, the channels of communication over which they passed messages, and the social organization of the ship.

This left unanswered the question: What IS inside the navigators? It did not seem likely to me that the PSSH provided a good account of what was happening in the minds of

the members of the navigation team. As I wrote up these navigation studies in *Cognition in the Wild* (Hutchins, 1995), I struggled to articulate how one might model the internal cognitive processes of the navigators. In chapter five of *Cognition in the Wild*, I argued that some sort of connectionist system was a better fit to individual cognitive processing than the PSS is. I designed and implemented some computational simulations that modeled individual cognitive function as connectionist constraint satisfaction networks. I constructed communities of such individuals to illustrate how the decision-making properties of a community as a whole could be changed by changing the patterns of interaction among the individual agents without changing their internal properties at all. This has implications for real-world activities such as jury decision making, for example.

Then, in chapter 7 of *Cognition in the Wild*, I addressed the individual cognitive function of learning in context. I created a clumsy thought experiment to explore how a written procedure might be learned if the learner had learning properties like those of a certain class of connectionist network. One network might learn the sequence of words in the procedure. Another network could learn the sequence of meanings of the words. Yet another network could learn the sequence of actions described by the meanings of the words in the context of the sensed local setting. These networks would each sense the world and pass constraints to the other networks.

At that time, many pieces of that puzzle were still missing. As far as I knew, the kinds of networks I needed did not exist. Little was known about interactions among multiple networks working simultaneously on a single problem. The system of networks and processes I imagined in the early 1990s was already multimodal, and in retrospect, I see that feature of the model as prescient. The sort of network I needed then now exists as a predictive processing system.

From the modern perspective, I can look back and see what was lacking in the model I imagined 30 years ago. It did not exploit a generative computational framework, nor did it maintain continuous coupling to its environment. Those are elements that have only become available in computational systems of the last two decades. It is only now, in the mid-2020s, that I think I am able to sketch an answer to the question of what a model of individual cognitive function should look like. And that is one aim of this paper.

## 2.3. Airline Operations

[3] From the late 1980s until 2016, my research village was the worldwide community of airline pilots. The cockpit of an airliner is a distributed cognition system, much like the bridge of a ship - except it is moving a lot faster. The papers titled "Distributed cognition in airline cockpit" (Hutchins & Klausen, 1986) and "How a cockpit remembers its speeds" (Hutchins, 1995b) are examples of the application of the distributed cognition approach to this domain.

**Figure 2.3. Briefing the approach to Christchurch, New Zealand in a Boeing 737**



*Photo by the author*

Through arrangements with NASA, Boeing, Airbus, and many airlines around the world, I made observations in the flight decks of airliners on hundreds of revenue flights. My research team and I interviewed pilots and took still photos of them at work in the cockpit. We made video recordings of activities in flight simulators. We investigated and contributed to the design of flight deck instrumentation, operating procedures, and

training programs. Employing (and enjoying) the participant observation aspect of traditional ethnography, I earned a commercial pilot license with qualifications in the first commercially successful airliner, the Douglas DC-3, and in two modern business jets. I also completed the training in the Boeing 747-400 (at Boeing in 1991) and the Airbus A320 (at America West Airlines in 1995).

Our investigations of flight crew activity are grounded in an ongoing long-term cognitive-ethnographic study of commercial aviation operations (Holder & Hutchins, 2001; Hutchins, 2007; Hutchins et al., 2006; Hutchins et al., 2009; Hutchins et al., 2013; Hutchins & Holder, 2000; Hutchins & Klausen, 1986; Hutchins & Nomura, 2011; Hutchins & Palen, 1997; Nomura et al., 2006; Palmer et al., 1993; Weibel et al., 2012). This ethnographic background allows us to interpret expert action and to ensure the ecological validity of our studies in high-fidelity flight simulators. There were many interesting objects of scrutiny in the domain of commercial aviation. My principal focus was the ways that flight crews understood and interacted with the automation in modern cockpits.

Over the course of the three decades I spent studying flight deck operations, methods and theory were changing rapidly. Methodological innovation included advances in the instrumentation of activity. The advent of inexpensive digital video brought a major change for all analysts of real-world activity and spurred the development of the behavioral science fields that focus on activity and interaction. Navigating in digital video is qualitatively different from using videotape. Using videotape, it is possible to compress time by moving fast-forward or -back, but time is still a continuous function of distance on the physical tape. Digital video provides wormholes through space-time. Moving from any temporal location in the recording to any other location is essentially instantaneous. Indexing is automatic, as every frame can be identified by its timestamp. Sensor technology has also advanced with motion capture and mobile physiology measures.

In the laboratory I shared with James Hollan at UCSD, we developed a Digital Ethnographer's Workbench. This included a set of digital tools to support the collection and analysis of field observations. For example, we created digital field notes with hyperlinks to digital scans of all the documents used by the flight crew.

Culturally elaborated real-world activity typically involves multiple operators in interaction with one another and with complex technical systems. Our object of study is

therefore complex, multiparty, multimodal, and socio-technical. Taking advantage of 70 years of development of sophisticated virtual reality environments, we instrument high-fidelity flight simulators and record the behavior of qualified flight crews in near-real-world situations. Modern sensor technology makes it possible to measure an unprecedented number of features of activity in such systems. However, the proliferation of measurements creates its own problems. Synchronizing and visualizing the relations among multiple data streams are difficult technical problems. Navigating rich data sets is difficult because of the sheer amount of data that must be managed. Data set navigation can be facilitated by good annotations and metadata, but providing even minimal annotation in the form of a timeline of events is a daunting and expensive task.

One of Jim Hollan's students, Adam Fouse, created an application called ChronoViz that provides solutions to many of the problems of measuring, analyzing, and visualizing the behavior of people in interaction, including airline flight crews (Hutchins et al., 2013). I will describe ChronoViz in more detail below in connection with the analysis of a few seconds of activity in a simulated Boeing 787 flight deck. At this point, let me just mention that ChronoViz supports the temporal alignment of multiple data streams. Timelines of key events can be generated and displayed within minutes of the conclusion of a simulator session. The entire data set can be navigated via any of the representations of any of the data streams. A great deal remains to be done, but we believe we have taken some important first steps toward using computational methods and good interface design to break through the analysis bottleneck created by the need to hand-code complex multimodal data sets.

## 2.4. Contributions of Cognitive Ethnography to Cognitive Science

In the new century, frameworks for understanding cognition changed too, with the rise of the fields I mentioned in the introduction that focus on the interaction of a person with their settings. Because no field or discipline has yet taken ownership of cultural-cognitive ecosystems, little is known about their function. Truly understanding how such systems work will require a large and sustained cognitive ethnographic endeavor.

# 3.   Related Fields

[4] Over the past three 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 the cognition of socio-technical systems so well, appears increasingly irrelevant to internal cognitive processes. As Clark put 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, 1998, p. 95)

Several approaches strive for an understanding of the nature of human cognition by taking seriously the fact that humans are biological creatures. All of these provide some useful conceptual tools for understanding real-world cognition.

Ecological psychology (Gibson, 1986) focuses on psychological phenomena as properties of animal--environment systems. To understand perception, one must understand the properties of the world to be perceived. To understand action, one must understand both the motor systems and their interactions with the world. A synergistic relationship grew up between ecological psychology and the development of the dynamical systems approach to cognition. The dynamicists emphasize that the system that matters is the brain, body, and world coupled in motion (Kelso, 1995; Port & van Gelder, 1995; Spivey, 2007; Thelen & Smith, 1994), while ecological psychologists have borrowed analysis tools from dynamical systems theory (Kugler & Turvey, 1987). Such accounts have successfully modeled many perceptual and motor processes. However, it is not clear whether high-level cognitive processes can be captured by more of the same kind

of process. The heterogeneous nature of real-world human action is a continuing challenge for dynamical system models.

Organism--environment dynamics become agent--environment interactions in embodied robotics (Beer, 2008). These efforts explore the ways that robotic agents can take advantage of structure in the environment to do thinking without representation (Brooks, 1991). For example, Brooks implemented simple robots that could autonomously navigate a setting. A low-level layer of control was designed to avoid objects. It is easy to build a robot that senses obstacles and turns away when one is encountered. This can be done simply by arranging the wiring that connects sensors to the motor system. No representations are needed. It is also easy to build a robot that can move to a particular distant visible target. Again, this can be done without representation. As Brooks says, "The second layer injected commands to the motor control part of the first layer, directing the robot towards the goal, but independently, the first layer would cause the robot to veer away from previously unseen obstacles. The second layer monitored the progress of the creature and sent updated motor commands, thus achieving its goal without being explicitly aware of obstacles, which had been handled by the lower level of control" (Brooks, 1991).

As the points of contact between organism and environment come to be seen as loci of essential processes rather than as barriers and boundaries to be crossed, the role of the body in thinking comes to the fore. These ideas are explored in two related contemporary approaches to cognition. In Europe, this is known as enaction. In North America, the embodied cognition perspective covers similar ground but from a different intellectual background.

The enaction perspective combines the philosophy of phenomenology (Dreyfus, 1982; Heidegger, 1962; Varela et al., 1991) with the cybernetic approach that appeared in the ecology of mind approach (Bateson, 1972; Dupuy, 2000). Building on the biological concept of autopoiesis (Maturana & Varela, 1987), the enaction perspective emphasizes that environments are not pre-given but are, in a fundamental sense, created by the activity of the organism (Havelange et al., 2003). The processes of life and those of cognition are tightly linked in this view (Thompson, 2007). 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.

Gibson's insight that perception is a form of action provided inspiration for a part of the philosophy of embodied mind movement (Hurley, 1998; Rowlands, 2006). For these authors, perceptual experience is grounded in regularities in the relations between sensation and action. These approaches view organism--environment relations in terms of coupling, coordination, emergence, and self-organization, rather than the transduction of information across a barrier. 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 (2002) introduced the idea of *sensorimotor contingencies*. In the activity of probing the world, we learn the structure of relationships between action and perception. These relationships capture the ways that sensory experience is contingent upon actions. Each sensory modality has a different and characteristic field of sensorimotor contingencies.

Embodiment is the premise that the particular bodies we have influence how we think. The rapidly growing literature in *embodied cognition* is summarized in Gibbs (2006) and Spivey (2007). Embodied cognition grounds high-level conceptual processes in bodily experiences (Barsalou, 2010; Calvo & Gomila, 2008; Johnson, 1987; Lakoff & Nuñez, 2000; Pfeifer & Bongard, 2007). One of the virtues of this approach is that emotion finds a natural connection to conceptualization through processes in the body. A subfield of embodied cognition examines the relations between gesture and thought (Goldin-Meadow, 2003; McNeill, 2005). Gesture studies highlight the coordination of talk with bodily action, demonstrating the multimodal nature of communication

Interactions between persons and their environments often simultaneously engage several modalities, speech and gesture, for example. It is now clear that inside the brain as well, the causal factors that explain the patterns seen in any one modality may lie partly in the patterns of other modalities. In fact, recent work suggests that activity in various cortical areas (e.g., visual and motor cortex, or visual and auditory cortex) unfolds in a complex system of mutual causality (Gallese & Lakoff, 2005; Sporns & Zwi, 2004; Wilson et al., 2004). Neuroscientists have thus become aware of the need to expand the boundaries of the unit of analysis to consider a wider cognitive ecology. In a review of psychophysiological methods Kutas and Federmeier say, "...the complexity problem presented by the mind--brain--body system may require new ways of thinking

about the kinds of measures we use and need to use because, in fact, the mind arises in a physical system that is distributed over space and time (Kutas & Federmeier, 1998).

Embodied cognition has its roots in psychology and is investigated mostly using experimental methods. The field of *embodied interaction* grows out of conversation analysis and linguistic pragmatics. It takes a more ethnographic approach to multimodality (Streeck et al., 2011). As the name implies, embodied interaction focuses on how people use their bodies in coordination with features of the social and material setting while interacting with one another. This is a bigger topic than the interactions of a single organism or person with a local environment, and it deals with concepts that are further from the organism-environment interface. Interactions among people are often carried out in public. To the extent that interaction can be read as a form of cognition, this is "cognition as public practice" (Streeck et al., 2011, p. 3). It is a form of cognition that can be recorded and analyzed.

The notion of semiotic resource is a core concept of the embodied interaction approach. This is nicely broader than the idea of external representation, although the so-called external representations can be seen as a subset of semiotic resources. A semiotic resource is an object, pattern, or event in the world that comes to have meaning for the participants in an activity/interaction. Words can be semiotic resources, of course, but so can gestures, facial expressions, and features of the material world. Researchers in this area speak of the mutual elaboration of semiotic resources. The phenomenon of environmentally coupled gesture is a prototypic case of this mutual elaboration. Goodwin analyzes an interaction between an archaeologist and a student who are examining discolorations in the dirt of an excavation.

"... by itself, the talk is incomplete both grammatically and, more crucially, with respect to the specification of what the addressee of the action is to attend to in order to accomplish a relevant next action. Similarly, the embodied pointing movements require the co-occurring talk to explicate the nature and relevance of what is being indicated. .... By itself each individual set of semiotic resources is partial and incomplete (Agha, 2007; Goodwin, 2007). However, when joined together in local contextures of action, diverse semiotic resources mutually elaborate each other to create a whole that is both

> greater than, and different from, any of its constituent parts (Goodwin, 2000)." (Streeck et al., 2011).

As more complex interactions are considered, the universe of semiotic resources grows so that, beyond mutual elaboration of two or three resources, we encounter the inter-elaboration of many semiotic resources, including those found in speech, body, activity, and setting.

Sequential organization is central to the way action is understood by participants. Conversational turns always build on what has come before. I imagine this to be a recursive process in which the meaning of utterance N depends on the meaning of utterance N-1, which in turn depends on the meaning of utterance N-2, and so on to the beginning of the conversation or to the limit of local memory. The process of constructing an interaction is dynamic. Speakers sometimes change the structure of a sentence while it is being constructed. Interactants can shift the meanings of their own and others' utterances by what they subsequently put in play, so that the meaning of utterance N may also depend on the meaning of utterance N+1.

The modalities in embodied interaction occupy a different level of description from the modalities of ecological psychology or dynamical systems. In embodied interaction, the modalities are gesture and talk, of course, but also environmentally coupled gesture, facial expression, body posture, how participants orient to one another, and how they position themselves in the setting and with respect to various semiotic resources.

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 gestures, are part of the activity in which concepts are formed (Alač & Hutchins, 2004; Gibbs, 2006; McNeill, 2005). That is, concepts are created and manipulated in culturally organized practices of moving and experiencing the body. Similarly, gesture can no longer be seen simply as an externalization of already formed internal structures. Ethnographic and experimental studies of gesture are converging on a view of gesture as the enactment of concepts (Goldin-Meadow, 2003; Núñez & Sweetser, 2010). 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 (Lakoff & Nuñez, 2000).

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 so-called "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.

Both embodiment and enaction stress the tight relation between thought and action (Alač & Hutchins, 2004; Hutchins, 2010b). Enaction shares with the dynamical systems approaches a commitment to circular rather than linear causality, self-organization, and the structural coupling of organism and environment.

Cognitive grammar (Langacker, 1987) and conceptual blending (Fauconnier & Turner, How We Think, 2002) are cognitive linguistic theories that describe phenomena that fit poorly with the theory underpinning mainstream cognitive science at the end of the 20th century. The predictive processing approach is much more congenial to these approaches.

Cultural Historical Activity Theory, with its emphasis on the social construction of thought, inspired other approaches that consider the cognitive consequences of social and cultural configurations (Daniels et al, 2007). Activity theory is the direct ancestor of the situated action perspective (Greeno & Moore, 1993; Lave, 1988; Lave & Wenger, 1991; Rogoff, 2003; Suchman, 1987). With its emphasis on the interconnections of developmental processes on all timescales (phylogenetic, cultural, ontogenetic, and micro-genetic), activity theory has been put to work by educational researchers as well (Greeno, 1998; Pea, 1996).

The rise of connectionism not only transformed theories of internal mental processes, but it also spawned a wider investigation of emergent phenomena at the supra-individual level. There is a growing literature on computational models of social and cultural systems. The emergence of language from interactions among agents is a particularly interesting area of research (Cangelosi & Parisi, 2002;Hazlehurst & Hutchins, 1998; Hurford et al., 1998; Hutchins & Hazlehurst, 1995, 2002; Hutchins & Johnson, 2009).

The field of collective intelligence focuses on the organizational principles that determine the cognitive properties of groups (Malone & Bernstein, 2015; Sunstein, 2007;

Surowiecki, 2004). Barbasi(2002) describes powerful regularities that explain how patterns of connectivity can change the cognitive properties of a network. The subtitle of Barbasi's book is "How everything is connected to everything else and what it means for science, business, and everyday life." While everything is connected to everything else, the patterns in the density of interconnectivity determine cognitive properties of the system, whether the system is an area of a brain or a group of governmental agencies responding to a crisis.

# 4. Framework

Guided by the findings of the related fields described in the previous section, I assume that every moment of human experience is multimodal, generative, and continuous (MGC). The theoretical instantiation of PP supports the generative aspect of this formulation.

In this section, I outline a theoretical framework that integrates these three aspects of cognition. I will use this framework as a guideline for creating descriptions of ongoing activity. It will guide our looking. The framework is multimodal in the sense that it includes all of the sensory and motor modalities. It is generative to capture the power of recent advances in generative AI to model some aspects of intelligent behavior. And it is continuously coupled to the sensorimotor surfaces, and via those surfaces, continuously coupled to the body and the world. The generative component is the heart of the model. This is imagined to be a connectionist network loosely modeled on a predictive processing system.

Predictive processing (PP) networks have three instantiations. There is a Theoretical Instantiation of PP, which is a description of an imaginary network of units. This theoretical instantiation is a computational framework that was inspired by a combination of a theory of physical thermodynamics (the Boltzmann machine) and by brain function. It was, as they say, 'neurally inspired', so, to the extent that it is accepted that the brain is a PP system (and Andy Clark's (2023) book provides ample evidence that this is so), a second instantiation of PP is an actual network of neurons implemented in meat. Let's call this the Biological Instantiation of PP. The theoretical instantiation of PP has also been given multiple implementations in silicon. These computational instantiations are what we now know as Generative Artificial Intelligence (genAI). Let's call these the Artificial Instantiations of PP. One of the challenges of

writing about this area is that in popular discourse, these instantiations are often conflated, and this conflation causes a good deal of confusion.

The very name Artificial Intelligence implies a claim to a particular relationship between the computational implementations and human intelligence. Most researchers in cognitive science take intelligence to be a property of the brain. This claim is strongest when the computation is implemented in neurally inspired networks. In both cases, genAI and the brain, the work is done by vectors of numbers. Both kinds of systems are composed of huge numbers of vectors that are very cleverly arranged and connected, constraining and responding to the constraints of one another. The vectors of numbers are implemented differently in neurons, in one case, and transistors in the other, but it's all just vectors of numbers in either case.

My interest is in theoretically hypothesized PP networks. It would be nice if theoretical PP networks were an accurate description of brain function, although, of course, now in the mid-2020s, they are still lacking many important aspects of nervous system operation. It is certain that no currently implemented computational version of PP is fully faithful to human cognitive function. Therefore, it is important to keep in mind that the PP component of the framework I describe below is neither the brain nor is it any currently implemented version of generative AI.

**Figure 4.1. A sketch of the MGC system**
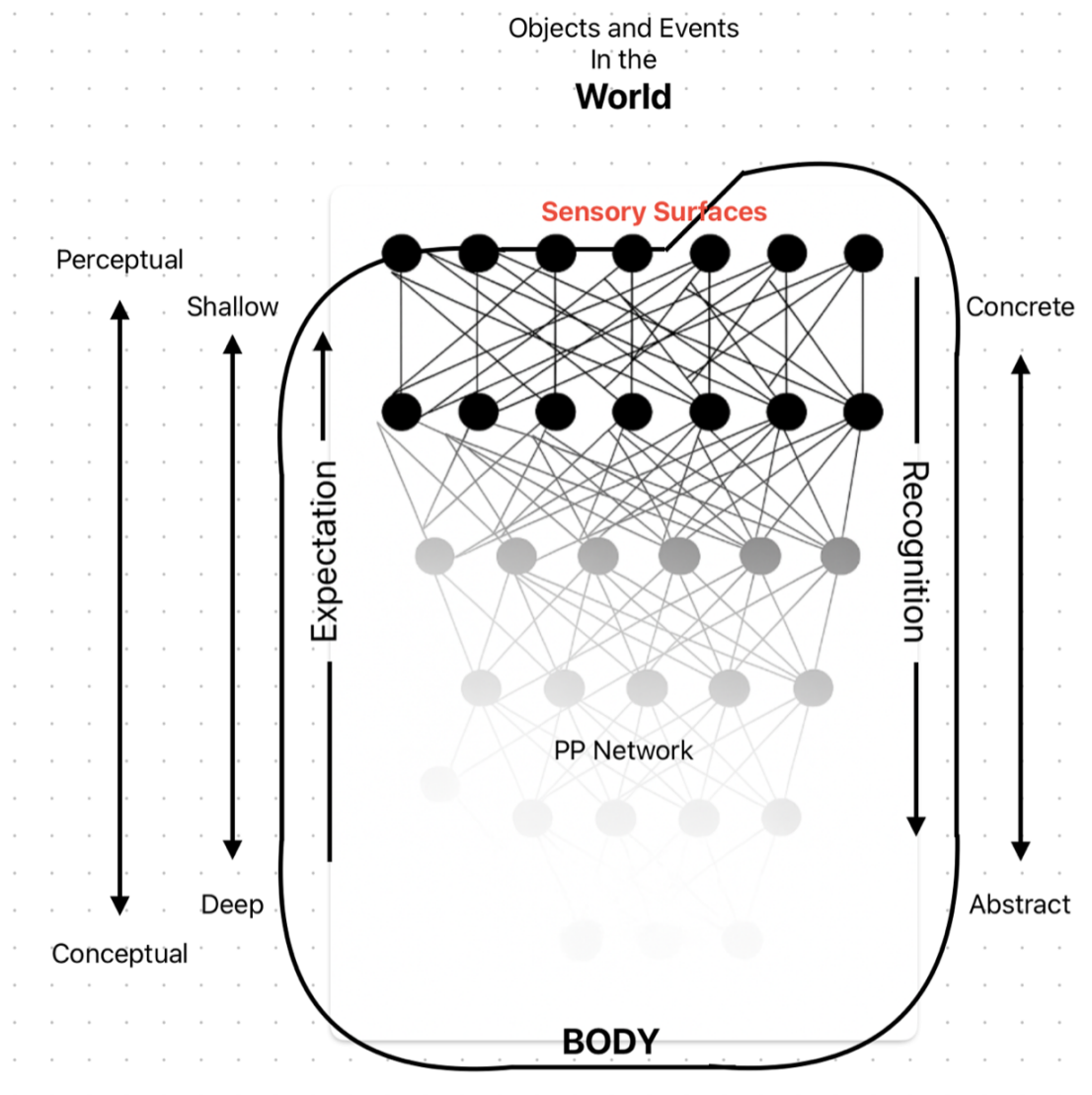
This is my sketch of a theoretical instantiation of the MGC system. It is not intended as a biological instantiation, although it borrows some concepts from neuroscience. It is not intended as a computational instantiation, although it borrows some concepts from generative AI. It is not strictly a predictive processing instantiation, although it borrows from that line of research as well, especially as it is presented by Andy Clark (2023). The diagram is grossly oversimplified, of course. The human brain is estimated to have around 90 billion neurons and has a complex architecture of regions.With this sketch, I only intend to suggest a few key functional relationships between the world, sensory-motor surfaces, and a very large neural network. It is these imagined relationships that guide our examination of real-world activity.

The body is in the world. The body contains a PP network. At the top of the PP network are layers that comprise the sensorimotor surfaces. These layers of the PP network are simultaneously in the network - shaped by their connections to layers below - and in the world - shaped by the physics of the world that impinges upon them. Below the layers that comprise the sensory surfaces is a deep connectionist network. The upper layers of the network that lie near the sensorimotor surfaces are associated with particular sensory or motor modalities covering both exteroception and interoception.These layers encode concrete features of the flow of sensation or action.

Deeper layers encode more complex features. For example, consider the representation of objects in various spatial frames of reference in the human visual system. There is a progression of spatial representations from shallow to deep as follows: Retino-centric - head-centric - body-centric - allocentric. Each successive representation in this sequence integrates more contextual information than its predecessor. The retino-centric frame of reference is modality specific. It answers the question, where is the object on the retina of the eye? The deeper representations appear in layers that have connections to other layers. They constrain, and are constrained by, the patterns of activation on other layers. The head-centric frame of reference, for example, requires coordination of the retino-centric representation with proprioceptive information about the orientation of the head. A body-centric representation occurs deeper in the network and requires the head-centric representation plus additional information from other sense modalities concerning the relationship of the head to the body. Allocentric spatial representations occur even deeper in the network as they require a representation of the space around the person in addition to representation in the other frames of reference.

Thus, while shallow layers encode features of the patterns of activation on the sensorimotor surfaces and are modality specific, deeper in the system are layers that span modalities. This is where the modes interact with one another. It is the realm of sensorimotor contingencies, hand-eye coordination, motor resonance, mirror neuron phenomena, and many more multimodal effects. Even deeper in the network are layers that encode abstract concepts.

Layers hold patterns of *activation* as internal structure. Layers are connected to one another by connective tissues that hold patterns of *influence* as internal structure. The connective tissues carry influence from part of the pattern in one layer to a part of the pattern in another layer. The pattern of activation on any given layer is determined by the patterns of *activation* on the layers to which it is connected and by the patterns

of *influence* in the tissue that connects them. Such a system can be described by configurations of activation across its layers. Inference (constraint from a shallower layer to a deeper layer) and generation (constraint from a deeper layer to a shallower layer) happen as the layers update their patterns of activation in accordance with the patterns of activation on the layers to which they are connected as mediated by the patterns of influence on the tissues that connect the layers.

The configuration of activation across all layers of the system at any given moment is known as the system state. As the system runs, the patterns of activation on the layers are updated. This means that the system transitions through states. This is how the relatively short-term changes of recognition and prediction happen. The entire inventory of possible states for the system is called its state space. For a given overall pattern of connective influence, each state will have a certain probability of occupancy. The operation of the system can be described as a trajectory or path through state space.

Learning in such a system happens by changing the patterns of influence on the tissues that connect layers. Changing these patterns of influence changes the state space of the system and changes the probability distribution across the states in the state space. What a generative network learns is to predict or reproduce its input. A PP network learns to predict the flow of sensorimotor experience. In order to predict this flow, a generative network must develop models of the processes in the environment that cause the flow of sensorimotor experience. This sounds incredible, but it works. I will have more to say about this in the section below discussing the Generative aspect (G) of the MGC system.

Notice the shallow/deep spatial metaphor. This depth dimension maps roughly onto content properties. Shallow layers encode specific, concrete features, while deep layers encode general, abstract concepts. In this diagram, sensorimotor surfaces are at the top, and deep layers of conceptual organization appear lower in the figure. This orientation reflects the grounding of cognition in action and perception rather than in abstract computation. Constraints that shallow layers exert on deeper layers implement inference, encoding, classification, and recognition. The constraints that deeper layers impose on shallower layers are the generative influences of decoding, expectation, and prediction.

This is an inversion of conventional network diagrams, which typically put abstract concepts and general category names at the top with sensorimotor processes at the

bottom. I have flipped the network for several reasons. First, it puts sensorimotor processes at the top of the diagram where they are perceptually salient for you, the reader. I want to do this to counter the implicit devaluation of sensorimotor processes in mainstream cognitive science. Second, it constructs a frame of reference in which layers near sensorimotor layers are 'shallow' and encode features, while conceptual layers are deep (rather than high). This will be useful to me later when I discuss the idea that language phenomena operate in 'shallow' symbols. Third, in this scheme, generative influence is oriented upward (that feels right), and inference is downward.

The Internal structure of the Imagined system is a complex, partially reconfigurable architecture. For our purposes, the details of this architecture will remain unspecified. The designers of generative AI systems know the general arrangement of the layers in the networks, but they do not know the details of the wiring (the learned connections among the layers of units) of the system after it has been trained. Neuroscience has made strides in mapping the architecture of the brain, but much remains to be learned.

There are limits on what we can know about the internal functioning of such networks beyond the fact that they learn to predict the statistical regularities of their environment. Suleyman and Bhaskar, describing contemporary generative AI systems said,

"In AI, the neural networks moving toward autonomy are, at present, not explainable. You can't walk someone through the decision-making process to explain precisely why an algorithm produced a specific prediction. Engineers can't peer beneath the hood and easily explain in granular detail what caused something to happen. GPT-4, AlphaGo and the rest are black boxes, their outputs and decisions based on opaque and impossibly intricate chains of minute signals." (Suleyman & Bhaskar, 2023, p. 149)

There is no way to know in detail how any particular prediction was created by either a brain or a generative AI system. When studying either the biological or artificial instantiation of the predictive processing system, it is possible to observe and document the regularities in the flow of experience that these systems must predict. It is even possible to determine where activation is highest under certain task demands, as is done in brain imaging studies. But it is not possible to know in detail how any particular prediction or action arose. In any case, I am developing a theoretical model and am not making a commitment to any particular artificial or biological instantiation of a network

system. Even without committing to details, it is possible to sketch a generic description that applies to most instantiations of generative networks but commits to none.

In this scheme, there is no separate downward and upward pass. The separation of passes is a feature of digital computational implementations. Instead, I imagine this theoretical system to operate through a single settling process in which constraints propagate continuously in all directions. The system simultaneously recognizes (downward) and generates (upward) the patterns of sensory activation. The entire network is in continuous resonance with the patterns of activity on the sensorimotor layers. The patterns of activation on the sensory surfaces are constrained simultaneously by the world, via the body, and by the dynamics of the entire deep network. I will have more to say about this continuous operation in the sections below.

# 4.1. Multimodal

The two perspectives on embodiment described in the related fields discussion produce two distinct kinds of multimodality. Each implies ways of speaking about modalities and the relationships among modalities as well as a typology of modalities. What counts as a mode in a multimodal system depends on whether one focuses on intra-individual function or on interactions among individuals.

The **embodied cognition** perspective and predictive processing build theory around the intra-individual functioning of a single organism. They focus on how sensorimotor processes are integrated into internal systems. In addition to the usually considered senses of vision, audition, touch, taste, and smell, we experience our own bodies via proprioception. We monitor the position and motion of our limbs. Our vestibular system keeps track of accelerations, including that of Earth's gravitational field. We sense how much muscle tension we apply to hold a position or to move. There are also what are known as interoceptive senses that monitor our internal bodily states, hunger, thirst, heart and respiration rates, blood pressure, pain, the states of our internal organs, and so on. Among our motor modalities are the control of skeletal muscles, voluntary and otherwise. Largely out of awareness, we also sense the body motion involved in speech, eye position and movement as well as the internal muscles that accomplish swallowing, and peristalsis.

In addition to sensory and motor modalities, we have deeper encodings of experience that integrate information from multiple sensory and motor modalities. The encoding of space is probably the best understood in terms of these deeper encodings. Spatial encodings integrate information from visual, auditory, and proprioceptive modalities. These may be egocentric, based on the own body or parts of the body such as the eye, head, and hand. Spatial coordinates may also have an allocentric (external) frame of reference.

Sensorimotor contingencies describe relations among modalities. But there are more subtle effects as well. 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, 2010; Gibbs, 2006, p. 61). Thus, we should expect that embodied, multimodal experiences are integrated such that the content of various modes affect one another.

In contrast to the embodied cognition approach, the **embodied interaction** perspective (Streeck et al., 2011) builds its theory around inter-individual processes, examining interactions among persons engaged in joint activity. This field identifies the communicative modalities: talk, gesture, facial expression, body posture, as well as features of the setting. It also explores super-segmental features of talk such as prosodic structure, tempo, rhythm, etc.

Most of the modalities identified in embodied interaction are already richly multimodal in the embodied cognition perspective. The production of speech involves motor, auditory, and conceptual processes. Other modalities may be involved as well, depending on the topic spoken about, as in the case of discussing the perception of odors or flavors, for example. The production of gesture involves motor, proprioceptive, and sometimes visual processes.

One of the relationships among modal contents highlighted by embodied interaction uses the phrase "mutual elaboration of semiotic resources." This is shorthand for something more complex that gives us an additional window into the relations between the modalities of embodied cognition (vision, audition, proprioception, etc.) and the modalities of embodied interaction. For some object or event in the environment to be a semiotic resource requires that its impression on the sensory surfaces be coupled to and predicted by a generative network. It is this coupling that produces the meaning of the semiotic resource. Some object or event is a semiotic resource when its impression on

the sensory surfaces is incorporated in a dynamic activation configuration such that it is **seen as** having a particular meaning. This dynamic activation configuration **is** the meaning of the semiotic resource. It is how the resource is seen as having a particular meaning.

When two semiotic resources mutually elaborate each other, the elaboration happens, not in the world of objects and events, but in the networks in which the semiotic resources are embedded. The networks pass each other activation. They have a relationship of mutual excitation. They may be entwined with each other, sharing some layers or segments of layers of network units.

Both the embodied cognition perspective and the embodied interaction perspective emphasize the fact that, while it is possible to separate modalities for the purposes of analysis, in fact, multiple modalities are almost always integrated in action. As we will see in the vignettes below, the relations among the contents of modalities, whether within or between persons, can be functionally important. Multiple modalities with congruent contents may mutually reinforce one another, providing stable representations. When the contents of the modalities are complementary rather than congruent, relations among modalities can be sources of variation in adaptive processes.

## 4.2. Generative

When it is read as a metaphor for processing in the brain, the generative system is coupled to a flow of sensory evidence, which it recognizes and predicts. There are many generative formalisms. I am not committed to anyone. All of the instantiations of predictive processing are composed of a deep hierarchy of layers of units with sensory and motor processes at the surface levels. Sensation projects downward through the layers of units while prediction projects upward toward the surface. Learning processes tune the connections among layers to bring the upward projecting predictions of each layer into agreement with the patterns of activation on the layer above.

The system learns the structure of the flow of sensation and uses what it has learned to predict not just the flow of sensation, but the elements of the learned structure that are implicated in the generation of the predictions of sensation. These are the so-called "models of the hidden causes of the flow of sensation." From the point of view of the predictive processing agent (or experiencing person), only the flow of sensation and the

models of its causes are available. The events and processes in the world that cause the flow of sensation are not observed by the agent; this is why they are said to be "hidden" from the agent. In learning to predict the flow of sensation, a predictive processing system must create internal processes that model or simulate the operation of these hidden causes.

In order to generate predictions that match sensation, the generative aspect of predictive processing must create a model of the workings of the world that produced the sensations. That is, a model of causes in the environment. The models of hidden causes are representations, but they do not represent the causes in any way that would be recognizable to an analyst or to the experiencing person. This is **representation without resemblance**. Nor is there any requirement that the models of hidden causes be correct, true, or veridical. They can be completely fictional as long as they efficiently predict the flow of sensation. The generative system is simultaneously active, predicting the contents of multiple modalities and doing so by simulating the hidden causes of those multimodal contents.

Clark (2023) distinguishes perception from sensation. Perception is sensation in the context of the network processes that predict it. Predictions change to better fit (reduce the difference from) the flow of sensory evidence. "To perceive is to find the predictions that best fit the sensory evidence. To act is to alter the world to bring it into line with some of those predictions" (Clark, 2023, pp. 212-213).

This generative capacity completes partial patterns, resolves ambiguities in sensed data, and imagines what "should be" even when the senses present something that should not be. As Andy Clark says, "There is a suggestive duality here such that to perceive the world (in this way) is to be able to imagine that world too - it is to be able to generate, using our inner resources alone, the kinds of neural response that would ensue were we in the presence of those states of affairs in the world" (Clark, 2023, p. 220).

This imaginative component is the source of confabulation (often incorrectly referred to as "hallucination") that is a cause for concern in generative AI chatbots. The chatbot may produce something that makes sense, given everything else it knows, but which is not true. The biological instantiation of PP is also subject to confabulation. Under some conditions, people may imagine familiar patterns when the senses carry no pattern at all. For example, Clark discusses a series of experiments in which primed subjects may report that they hear the melody of the song *White Christmas* when they are presented

with pure noise (Clark, 2023, p. 23). I was tempted to say that people can imagine familiar patterns when the senses carry no discernible pattern. But, of course, the point here is that when a subject hears *White Christmas* in pure noise, they *have* discerned a pattern. It just happens to be a pattern that did not exist in the sense data. This is a simple illustration of the fact that discernment is an active generative process.

An important feature of the generative capacity of predictive processing is that it provides a natural explanation for the fact that people routinely experience useful patterns that go beyond what is suggested by the sensory evidence. For example, in the presence of a spatial array, we may perceive the array, and perceive additional culturally supplied structure superimposed upon the array. We can see a bunch of stars as a constellation with a definite shape. We may even imagine lines in the sky tracing a connection between the stars of the constellation. This "seeing as" is an essential cognitive process that is the basis of a large family of cultural practices. We can see a line of people as a queue. We can see a line drawn on a navigation chart as a projected ship's track. We can see an array of numbers as a sequence or a scale of numbers. The generative capacity allows us to project internally generated structure onto sensation to produce experience that has emergent properties that are not present in either the sensation or the projected structure alone. In concert with culturally constructed environments for action, this projection of structure is an incredibly powerful ability. As I will show in the vignettes below, it is a way to produce what have traditionally been known as high-level cognitive processes while deploying low-level perceptual processes in culturally organized settings.

With experience, expectations become structured. The expectation generator learns from experience to predict sensory evidence. "What matters for our purposes is just that the generative model - however installed - is a learnable resource that will enable a system to self-generate plausible new versions of the kinds of data seen in training" (Clark, 2023, p. 220). The generative system must produce predictions on multiple time scales and at different levels of specificity. Because expectations can appear at any depth of the network, they can be general or abstract so that a range of details can be accommodated without violating the expectation.

Active inference is the term introduced by Karl Friston and colleagues as a way of highlighting the unity of perception and action under schemes in which perception aims to find the predictions that best fit the world, while action aims to make the world (starting with simple bodily motions) fit the predictions. One imagines or predicts the

movements that will accomplish some goal, and those imagined movements become a motor plan.

Action has another role here. Sometimes a generative model that includes the prediction of the sensory consequences of own action can improve the prediction of sensations that are caused by other processes in the environment. For example, understanding the gestures of others in interaction involves not only visual processes, but motor processes as well through the phenomena of motor resonance. When we see others acting in the world, our own motor systems may activate. The prediction of the visual sensations of observing someone else making an action may be improved by activating the motor circuits involved in making that action oneself. In these cases, visual and motor processes are coupled. Seeing the other person gesture activates one's own motor system, which shapes what one sees. In this way, the multimodal operation of predictive processing implies or subsumes mirror neuron phenomena. This is an unintentional communication of modal contents, a sort of contagion of motor system activation, carried by vision.

As Campbell and Cunnington say about the brain interpreted as a PP system, "Firing during action observation is not merely driven by visual input, rather it constitutes a part of a generative model actively predicting sensory input." (Campbell & Cunnington, 2017, p. 196) "The association between observed and executed actions built through common experience leads to the sensory input of observing another's action feeding forward as motor representations, then priming a matching motor plan." (Campbell & Cunnington, 2017, p. 198). Parts, but not all, of the activation patterns generated in predicting the sensory consequences of one's own action match and reinforce the activation patterns generated in predicting the sensory consequences of observing the other's action.

A generative model that includes own action (or sub-threshold activation of one's own motor plans) provides a more robust and more accurate prediction of visual sensation than processing visual input alone. This continuous activation of all parts of the generative system across modalities and up and down the hierarchy of network levels is an important feature of the MGC system. Note that while PP implies the existence of mirroring phenomena, the imitative aspect of mirroring, that received so much attention early on, is a side effect of the activation in support of accurate prediction of sensory input in observation. In this sense, the predictive processing framework implies the mirror neuron system, not as the main event, but as a component of a much larger

simulation in which the models of hidden causes of sensation are hypotheses about the operation of the experienced world. Mirror neurons do not exist primarily in order to imitate.

The larger generative system that subsumes mirror neurons will enter the discussion below in connection with motor resonance in teamwork and with the role that cross-modality representations may play in joint actions that are complementary rather than imitative.

The three instantiations of predictive processing networks share an important property. Each is composed of a large number of similar units. The internal world of a generative AI system, the artificial instantiation, is composed of a large number of computational units of just a few types. The internal world of a brain, the biological instantiation, is composed of very large numbers of neurons. There aremore types of neurons in a brain than there are types of units in an AI system, but even a brain is a relatively homogeneous system. The theoretical instantiation of PP systems relies on thehomogeneity of the component parts to make possible the formal mathematical models of its operation.

Somehow, such homogeneous collectives, in brains and generative AI systems, learn to model the heterogeneous constellation of objects and events "out there" in the world on the other side of their sensory surfaces. These systems **model a heterogeneous world through a heterogeneity of dynamics on a homogeneous substrate**.

## 4.3. Continuous

I first caught a clear glimpse of truly continuous organism-environment coupling in the writings of Maturana and Varela on autopoiesis and cognition (Maturana & Varela, 1987). This property also features prominently in the work of the intellectual descendants of Maturana and Varela, in the enaction approach, and in dynamical systems models of agent environment interactions. By **continuous**, I mean that the mind is always on. Sensory systems continuously conform to and track sensory evidence. The MGC system is always operating, shifting, and learning. It is always acting to shape, change, and control bodily sensation. Many of these processes are unconscious. Prediction and recognition are continuous and simultaneous.

According to the ecological psychology, dynamical systems, and ecology of mind approaches, control of sensation through action is part of perception (active sensing). That means that the unit of analysis must include the world sensed and acted upon. Predictive processing provides a model of a single system centered on the sensorimotor points of contact between organism and environments. The continuous nature of the organism-environment coupling implies that the environment must be an element of the ongoing process. The environment is not simply read and then discarded while processing is carried out on a representation of it; rather, internal processing is in continuous interaction with and is continuously shaped by the environment. The state of the network is always resonating with the activity on the sensory surfaces.

Of course, the structure of an environment can be learned, and once learned, a PP network can generate both the sensation that would appear were the structure present and the hidden models of the causes of that sensation. In that sense, an environment can also be processed while not actually present, as in the case of imagination, dreaming, and other forms of what has traditionally been called 'offline' processing. But some environment is always present. And unless the continuous connection is suppressed, that environment will participate in the dynamics of the MGC system processing.

In such a scheme, there is no input and no output. The concepts of input and output are analytical conveniences. They are conceptual devices that make it easier for researchers to hold parts of a multipart concept steady while other parts are manipulated. However, input and output should not be mistaken for features of the system we are describing.

Similarly, the continuous process of generation of prediction and comparison of prediction to sensory evidence is not best conceived of as a loop. Several of the contributing fields conceive of the organism-environment connection as a loop or invoke circular causality. Bateson (1972) and Clark (2001) speak of information traveling in a loop from the environment to the person and back out into the environment. Showing that such loops reach out beyond the skin and skull of the person was a major contribution to the development of the extended cognition perspective. The loop is true to the phenomena in the sense that it holds the correct pieces together. It is an advance over conceptions that "cut any of these pathways in ways which leave things inexplicable" (Bateson, 1972, p. 459). However, a loop implies a sequence of events. In a loop, one thing is followed by some other thing, which is followed by some other thing, until the first thing is again visited. The sequential character of the operation of the loop is a feature of the trajectory of attention of the analyst to one element of a

system, then to another, and so on. This renders a complex system understandable to analysts, but it is not true to the continuous character of the coupling between organism and environment.

How can we grasp the nature of this unified coupled system? Are there any useful metaphors? Consider musical instruments. Most musical instruments take an input (a key press, a string pluck, a puff of breath) and produce a note in response. The theremin is different (See the theremin video at [Theremin Video](#)). Unlike other instruments, when it is ON, the theremin is always vibrating. The theremin is a tone generator in which the tone is responsive to the electrical currents in wire antennas that protrude from the instrument. To play the instrument, one moves one's hands in the electrical fields surrounding the wires. Note that the theremin is coupled to its environment and produces a tone even when no hands are present. Placing one's hands near the wires changes the nature of the current in the wires, which is a key element of producing the tone. Moving one's hands produces continuous changes in the current in the wires and produces continuous changes in the tone. The theremin is thus continuously coupled to, constrained by, and responding to its environment. There is no sequence of causes. The tone is produced by the interaction of the tone generator with the characteristics of the electrical fields around the antennas. The generation of the tone does not precede the electrical fields, and changes in the fields do not precede the changes in the tone. The tone produced by the theremin is the audible manifestation of the ongoing continuous coupling of the tone generator to the electrical fields surrounding its sensing wires, just as perception in the MGC system is the manifestation of the ongoing continuous coupling of prediction, sensation, and the environment. The MGC system is always resonating in all modalities, where each modality can be thought of as generating a tone that is continuously constructed by the meeting, at its sensory surface, of its internal models with the sensory consequences of its surroundings.

It is difficult to conceptualize the entire system of coupled internal and external ecosystems as a single dynamic entity. The generative aspect of predictive processing is crucial. Prediction, sensation, perception: None of these processes precedes or follows another. They are all simultaneous.

It is unfortunate that the word prediction has a temporal sequence built into its morphology, 'pre - diction'. The fact that this approach does not have a more appropriate name is an indication of the lack of fit between the phenomena and the current inventory of conceptual tools in the behavioral sciences. As analysts, we must take care

to avoid mistaking features of our own clunky stepwise thinking for properties of the system we seek to understand.

## 4.4. Setting the stage for MGC and language

In the discussion of the role of the PSSH in the socio-technical system of ship navigation, I argued that some sort of connectionist system was a better fit to individual cognitive processing than the PSS is. However, since humans are clearly avid symbol processors, this stance raises the question, where ARE the symbols that humans and genAI systems so avidly process? Symbols are clearly processed in the socio-technical system, but in what ways are symbols processed in individual cognitive function?

Forty years ago, Rumelhart et al. (1986) noted that people are good at three things: 1) recognizing patterns, 2) manipulating the physical world, and 3) imagining simple dynamical processes. Recognizing patterns and imagining simple dynamics are nicely accounted for by the generative capacities of PP systems. Manipulating the physical world is accommodated by the active inference interpretation of motor processes.

Humans inhabit a world of symbols. Clark (1998) described surrogate worlds, systems of symbols with which people interact. Rumelhart et al. (1986) addressed this situation in an account of doing place-value multiplication with paper and pencil.

> "Each cycle of this operation involves first creating a representation through manipulation of the environment, then a processing of the (actual physical) representation by means of our well-tuned perceptual apparatus leading to further modification of this representation. By doing this we reduce a very abstract conceptual problem to a series of operations that are very concrete and at which we can become very good. ... This is real symbol processing and, we are beginning to think, the primary symbol processing that we are able to do." (Rumelhart et al., 1986, pp. 45-46)

As the MGC framework implies, with experience, people may learn to imagine the dynamics of these external symbolic worlds. Rumelhart et al. go on to say, "Not only

can we manipulate the physical environment and then process it, we can also learn to internalize the representations we create, 'imagine' them, and then process these representations -- just as if they were external." At the time they were writing, Rumelhart and his colleagues did not have a computational account that could do the things they describe. They had no coherent mechanism to internalize, represent, and imagine external states of affairs. The generative component of MGC provides exactly the capabilities called for here. However, the meanings of the key terms 'internalize', 'represent', 'processing', and 'imagine' are all transformed by the shift to MGC. Under MGC, representation is without resemblance between the thing represented and the patterns of activation that represent that thing. Internalization is no longer conceived as a transport of a pattern across a boundary. It is, rather, a process of coordination and coupling between external and internal events. Processing is the dynamic interaction of patterns of activation in the generative system. Imagination is the generation of expectation.

Earlier, I described an MGC system in continuous interaction with a world. Let us now consider what happens when there are symbols in that world to which the MGC system is coupled.

This is an important class of phenomena for many reasons. Language - composed of strings of symbols - is the most complex and arguably the most important aspect of culture. Understanding the implications of the presence of symbols reshapes our understanding of how symbols mean what they mean (symbol grounding), and it bears on the question of the origins of symbolic representation (which I will characterize as the ungrounding of sign forms). The MGC system in interaction with a world of experience that includes both physical symbol forms and the objects and events to which symbols refer has different properties than an MGC system in a world without symbols, and these differences transform our understanding of culturally mediated learning. Determining the role of symbols in the MGC environment is essential to understanding how communities of MGC systems are woven together in the fabric of culture.

The key concept here is that **symbols are shallow**. Like all other phenomena of the world surrounding the MGC agent, physical symbolic forms penetrate the human cognitive processor no deeper than the sensory surfaces. The generative processors are deep, and they represent the following elements - without resemblance - as they appear in the flow of sensation: the symbols themselves (semiotic resources in whatever

physical form), the relations among the symbols, and the relations between symbols and the things to which the symbols refer. People can, of course, imagine the dynamics of familiar symbolic systems, including language, in the absence of the physical symbols. Reasoning **is** discourse, whether it appears in interaction with others or in inner speech. The process of generating the predictions of the flow of words drives the deeper layers of the network through the meanings that constitute thought. Yes, thought is mediated by language, but it is not conducted in the manipulation of symbols. Physical symbol forms are represented only in shallow layers of the network. Thought is conducted in the passing of activation among the deeper layers of the network, where the meanings of symbols reside. The deeper layers contain representations that generate the patterns of activation on the sensorimotor surfaces that match the patterns that would appear were the physical symbols present in the world. In this scheme, perception, meaning making, and reasoning collapse into a single adaptive process.

From the point of view of the MGC system, the causes of symbolic sensation are "out there," hidden in the world on the other side of the sensory surfaces. Scientists infer that MGC systems learn models of these hidden causes of sensation. The systems must do so in order to make accurate predictions. The fact that both humans and several classes of generative AI systems, including those known as Large Language Models (LLMs), can predict both the syntax and the semantics of natural language sentences is evidence for the existence of models of the hidden causes of language structure.

Such models have never been directly observed in any instantiation of predictive processing. Not in a brain, not in any generative artificial intelligence system, and not in any hypothetical MGC. These models are inferred to be present, but like the predictions mentioned by Suleyman and Bhaskar (2023) above, their operation plays out in "opaque and impossibly intricate chains of minute signals."

Cultural models provide additional evidence for the existence of models of hidden causes of sensation. There is a literature on the mental models or cultural models that people use to make sense of the deeper structure of language and thought (Gentner & Stevens, 1983; Holland & Quinn, 1987). Psychologists, linguists, and anthropologists have been able to identify many underlying models of the operation of the world. Any good LLM chatbot can now be interviewed as an anthropological informant to demonstrate that it has acquired, from its exposure to massive amounts of text data alone, the same models that were painstakingly abstracted from discourse and text documents by researchers in the 1980s. In the 1980s, it was believed by most

researchers that these models were represented in the mind by symbolic propositions. Today's chatbots demonstrate mastery of these mental models without relying on internal symbolic propositions.

There is a two-way relationship between this hypothetical MGC model and the observations of ongoing real-world activity described below. I pull from the model conceptual structures that help me understand observed activity. These elements of the model underwrite my interpretations and analyses of what is happening when people act in their everyday lives. The MGC framework helps me judge the plausibility of phenomena that appear in the data. I push back into the model those phenomena I encounter that meet two criteria. First, they are observed but are not yet part of the functioning of the model. Second, while they are not yet part of the model, they fit with the other terms of the model. I use the MGC model as a framework for the interpretation of activity on the one hand, and as a receptacle for phenomena that are newly observed on the other. One of my long-term goals is to push descriptions of some phenomena right through the enriched MGC model and onto the research agendas of those who investigate all three instantiations of PP networks, theoretical, biological, and artificial.

## 4.5. The control of entropy

Brains control the uncertainty of experience in one of two ways: by changing (tuning/learning) a model that produces predictions that match the flow of sensation or by acting to change the world (or the organism's relationship to the world) to bring the sensed world into alignment with that which is expected or predicted. Both changing the model and changing the world can improve the match between predictions and sensation. The brain is in the business of making experience more predictable. Learning from experience reduces the uncertainty of experience. This is true of all animal nervous systems (autopoiesis). The big advantage for humans over other animals is culture. Most of human experience is culturally organized, and culture makes the world to be experienced more predictable. One could say that the brain and culture are in the same business; the business of controlling the uncertainty of experience. Culture controls uncertainty through processes operating via different mechanisms and on different time scales from those in the brain.

## 4.6. Imagine a world

This sketch of the MGC framework leaves many questions about human cognition unanswered. Applying it to observations of real-world activity raises additional questions. The knowledge or skills required to answer many of these questions are beyond my abilities. I'm hopeful that those who do have those skills will take up some of the questions raised here.

This lens for viewing cognition does not replace or deny other descriptive frameworks. It mostly complements them, showing us phenomena that go unseen or unnoticed from other perspectives. Occasionally, we will see that this view contradicts or undermines claims made based on other frameworks. The question is not "is this framework true?" Rather, it is, "is this framework useful?" The most important question is: "How does changing the model of human cognitive processing change the way we understand observable behavior?"

Now, please join me in imagining communities of people, each understood to be an MGC system, inhabiting historically contingent cultural cognitive ecosystems. I will attempt to take the MGC framework into the wild to interpret the cognitive aspects of everyday activities.

# 5. Ship Navigation Vignettes

In the early 1980s, while I was working on ship navigation, I received an invitation to a conference on mathematical reasoning. The conference organizers assumed that my work with navigators would give me insight into the nature of mathematical reasoning. They asked me to answer the question: What sort of mathematical reasoning do navigators do? I declined the invitation because the answer to their question is: Essentially none.

This is not to say that navigators do not perform complex computations. For example, drawing a line of position is a way to compute all locations that lie on a given bearing from a given landmark. In a simple Cartesian space, one could write an equation of a line in the form $(y = mx+b)$ to do that computation. In such a setting, one might imagine navigators reasoning about slopes and intercepts. Of course, navigators do no such thing. The quartermasters who perform the navigation functions on the bridge of a ship typically have only a high-school education with no training in math beyond

arithmetic. And the problems they actually face are much more sophisticated than that. The surface of the earth is not a plane, and while the locations in the space depicted on a navigation chart can be specified with two coordinates (latitude and longitude), the equations that define a given line of position are complex and involve non-Euclidean geometry.

Ship navigation is a special domain of activity because the material setting of the navigation bridge is the product of a centuries-long cultural evolution. It includes a rich collection of tools, some of which incorporate representational formats that are thousands of years old. For example, the Mercator projection is hundreds of years old, while the 360° direction circle is thousands of years old, and the idea of cardinal directions is so old that its origin predates written records.

Navigators are not expected to innovate. They are required to perform a small repertoire of well-understood procedures, each of which is tuned to the solution of a very limited set of problems. They carry out complex computations through the physical manipulation of common navigation tools. For example, the problem of computing all the points that lie on a given bearing from a given landmark is simplified by the Mercator projection chart. The Mercator projection chart is not just an accurate map of an area; it is a computational medium in which straight lines have the property of being lines of constant bearing. This means that the simplest of plotting tools, a straightedge, placed on a Mercator projection chart, defines a set of points that have an essential property.

So, how is it possible for navigators to perform complex computations without doing any math reasoning? The short answer is that they rely on a combination of the properties of the material setting and embodied domain knowledge. The longer answer will become visible in the application of the MGC approach to the navigators' ongoing activities in the sections below.

The MGC framework implies that cognitive processes should be visible in the fine details of the engagement of a whole person with a whole culturally organized world. In the following sections, I will attempt to perform such an analysis.

The first four activities analyzed in this section are drawn from observations of the navigation team on the bridge of a large US Navy ship as it enters the narrow channel at the mouth of its home port. The USS Palau[5] is an amphibious assault carrier. It is

about 200 meters long and looks like an aircraft carrier, having a flat flight deck and an "island" tower protruding five levels above the flight deck.

The navigation bridge is on the second level above the flight deck. There, the captain, the officer of the deck, the navigator (officer), the keeper of the deck log, helmsman, engine order telegraph operator, and bosun's mate stand watch during the entry to harbor. Also on the navigation bridge are two members of the navigation team, a plotter and a bearing recorder.

The Plotter, P, is a senior quartermaster chief. He is an experienced navigator and is formally responsible for the quality of the work of the navigation team. At the plotter's side is the Bearing Recorder, BR. BR is a quartermaster second class with just a few years of service in the navy. He has not much experience as a navigator and is an apprentice to the plotter. BR is in contact via sound powered phone with two other members of the navigation team, the bearing takers, who are stationed outside the bridge on the wings of the ship. Out on the wings, the bearing takers each have a telescope with a built-in compass that allows them to determine the bearing from the ship to various landmarks along the route of the ship.

P and BR stand side by side at the chart table, which faces starboard on the navigation bridge. A chart of the harbor is on the chart table. Lying on the chart are plotting tools, a pencil, and BR's wristwatch. BR has the bearing logbook and an ink pen on the chart in front of him.

The ethnographer stands at the aft edge of the chart table observing the crew. There is a video camera aimed at the chart area and mounted in the overhead above the ethnographer. P wears a lapel microphone recording to one channel of a stereo recording system. A microphone above the chart captures ambient sounds on the other channel of the recorder.

The navigation chart is a key tool in the position fixing activity. Weeks before this entry to harbor, P modified the chart, adding two features that are not on the chart as published. Throughout the navy, some short distances are expressed in units of yards. P has added a hand-drawn yard scale to the bottom of the chart. He has also planned the entry of this particular ship to the harbor, laying out in ink a projected entry track. The track consists of a series of straight-line segments laid out on the right side of the navigation channel. Each transition from segment to segment requires a turn, and the

entry track is marked such that each turn is preceded by a scale showing yards remaining to each next turn point.

Other tools appear in the activity: a pencil, the hoey, and dividers.

**Figure 5.1. Plotting tools: Hoey (left) and dividers (right)**
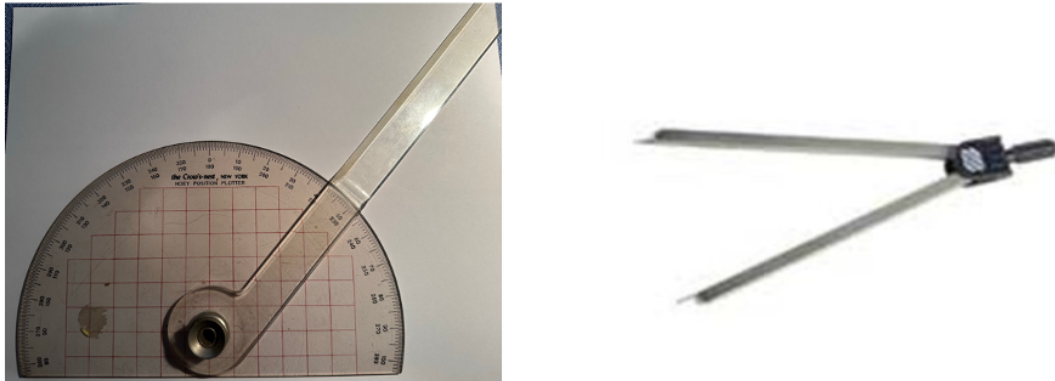


The hoey is a plotting protractor. The hoey base is a semi-circle marked with degree scales. There are 360 degrees in a full circle. A plotting arm rotates around the center of the degree circle. The arm is attached to the base with a lock ring screw that can be tightened and loosened to adjust the amount of friction between the arm and the base. The rectangular grid of red lines on the transparent hoey base provides for alignment of the hoey base with the chart. Laying one of the vertical red lines on a charted line of longitude or laying one of the horizontal red lines on a charted line of latitude brings the hoey base into alignment with the directional frame (relative to true north) of the chart. The dividers (confusingly also known as a compass) are a two-legged tool with adjustable friction to hold the angle between the legs and thus the span between leg tips.

The position of the ship is determined or "fixed" in an activity called the "fix cycle."

The steps in the fix cycle are:

1. Plot the ship's position as the intersection of three lines of position. Each LOP is defined by a bearing from the ship to a landmark - a visually prominent feature along the sides of the navigation channel.
2. Compute the ship's speed (optional, if there is reason to believe it is changing).
3. Estimate where the ship will be at the time of the next position fix. Draw this on the chart as the estimated position EP.
4. Choose three landmarks to use for plotting the ship's position at the next scheduled fix time.
5. At the scheduled fix time, observe and record the bearings to the chosen landmarks.

We will look at just three elements of the fix cycle: Plotting a single LOP, computing the ship's speed from distance traveled and time elapsed between fixes, and choosing the three landmarks to use to construct a future position fix.

On a cool spring afternoon in 1984, the USS Palau returned to its home port, San Diego Harbor, following several days of exercises in the waters off Southern California. Standing at the chart table on the navigation bridge, the plotter P and the bearing recorder BR were using visual bearings on familiar landmarks to plot the ship's position every three minutes. As the ship approached the harbor entrance, the navigation team plotted a position. P measured the speed of the ship and projected an estimated position where he expected the ship to be at the time of the next position fix. P and BR then chose three landmarks to use in plotting the ship's next position.

# 5.1. Plotting a Line of Position

[6] Once commenced, the cycle of activity for plotting a fix is a true cycle with no beginning point. For this analysis, I take as the starting configuration P and BR standing side-by-side at the chart table. BR hears the report of the bearing to the first landmark and reads it back to the bearing taker on the wings. This read-back also makes the bearing available to P, who has the hoey in his hands on the surface of the chart in front of his body.

Each vignette consists of three parts. 1) A description of observed activity. This description appears in *italic* font. 2) A richer re-description of the observations, including background processes and the history of the activity. 3) A discussion of the cognitive import of the observed activity as revealed by the MGC analysis. Let's put on our MGC goggles and take a closer look at the fifteen seconds of activity involved in plotting a single line of position.

This vignette begins with a moment of distraction. The keeper of the deck log DL, asks P if the ship has passed a particular channel marker buoy. Just as DL finishes his *question, BR reports the first bearing, "057 Hotel del." P replies quickly to DL, "Wasn't* watching." Then P turns to BR and says, "Huh?" to prompt BR to repeat the bearing, which he does, saying, "Zero five seven." But without specifying the landmark. P says, "'kay" and proceeds to plot the LOP.

## 5.1.1. Locating the bearing on the Hoey

Here (in *italic* font) is a description of the fine detail of P's multimodal engagement with the plotting task.

*BR reports the bearing to the landmark called Hotel Del (Coronado) as zero five seven degrees.*

*P leans over the chart and hoey. The hoey base is not aligned with the chart but rotated about 30 degrees counterclockwise. This orientation of the hoey puts the scale value 057 above the plotting arm between the plotter's hands and in clear view.*

*P's left hand rests on the hoey base, pinching the hoey lock ring between the side of his index finger and the pad of his thumb. The other fingers on P's left hand are folded under his hand, and he holds a pen between his thumb and the edge of his palm. The heel of his left hand rests on the surface of the chart. The fingertips of his right hand rest on the hoey arm.*

*Presently, the plotting arm lies across a scale value that is greater than the bearing of the current landmark.*

**Figure 5.1.1. P locating bearing on hoey**

*Frame from a video by the author*

Let's examine P's experience of this moment, considering only the proprioceptive, visual, and tactile modalities.

The plotter's proprioceptive experience includes his arms providing some support to his inclined torso, as well as the location of his hands.

His visual experience, with eyes focused on his hands and the hoey, is entirely congruent with this aspect of his proprioceptive experience. He can see his hands and arms in contact with the chart and the hoey. Together, proprioceptive and visual modalities feed deeper representations of location in body-space and local space. The way the view of the hoey changes with changing his neck, head, and eye angles to maintain gaze on the hoey base as he leans further forward is predictable from the

sensorimotor contingencies of vision. Changes in proprioception are consistent with changes in visual experience. The distance from eyes to chart may also be sensed in the vergence angle and focus depth of the eyes.

P's tactile sense adds information that is not available to proprioception or vision. He can feel the smoothness of the chart and the even smoother plastic surface of the hoey. Perhaps he feels that the hoey is cool to the touch. The fingertips of his left hand experience the roughness of the surface of the knurled nut that is the hoey lock ring. The sensorimotor contingencies here are interesting. When the fingers are stationary in light contact with the knurled surface, there is little sensation of roughness. Putting torque on the ring by applying lateral shear (experienced haptically) at the point of contact between the finger and the knurled surface adds to the tactile sensation of roughness. Knurls in the ring produce a prominent sensation at the fingertips when under tension. The muscle tension required to turn the ring is congruent with the perceived prominence of the knurls. Dragging the fingertips over the knurled surface maximizes the sensation of roughness.

The protractor scale is a domesticated space. It is a cultural construct that exhibits a number of cultural conventions. A circle is divided into 360 equal-angular increments, each conventionally known as a degree. This convention ties the hoey protractor to the basis of direction and location (latitude and longitude) frames in navigation as well as to the expression of the magnitude of angles. Scale values increase in the clockwise direction. This convention links this activity to circular scales and gauges of all sorts. There are large tic marks every five degrees. Ten-degree increments have large tic marks with numeric labels. There are medium size tic marks every full degree and small tic marks at half degrees.

Finding any scale value that is not a multiple of 10 requires interpolation. Interpolation requires counting and a recognition that exploits seeing a directional trajectory across values on a scale. The directional trajectory is created by the expectation that moving attention clockwise will bring larger values into focus and moving attention counterclockwise will bring smaller values into focus. The trajectories are built into the body of the practitioner in the habits of attending. Just as the knowledge of rotating a screw clockwise to tighten is built into muscle memory, culture is inscribed in human bodies.

When we regard this stretch of activity using the MGC framework, what do we see?

First, the plotter engages his world through **multiple modalities**. He engages this activity in visual, proprioceptive, and tactile modalities (among others).

Second, he is **continuously** and simultaneously sensing, expecting, and perceiving in all of these modalities. It is important to stress the simultaneity and continuity of the process. Not only are all of the senses simultaneously active, the generation of expectations, the registration of sensations, and the merging of expectation with sensation, which produces perception, are continuous processes. In this activity, the environment is always present in the plotter's thought.

Third, human experience in consequential activities is thoroughly **cultural**. The protractor is a convergence of several cultural conventions. A number sequence (a culturally conventional abstraction) over a conventional range (0 -- 359) is conceived as a spatial sequence in which each value has a specific location. This combination of number sequence with spatial sequence creates a number line. Wrapping the number line into a circle makes it possible to see the differences between numbers as angles. These abstract angles could be conceived as numeric quantities by applying arithmetic functions to the numbers. Or they could be conceived as shapes by imagining the span between locations on a circular number line. Of course, the fact that the number line is wrapped around the circle in a clockwise direction is also conventional.

This example illustrates an important general phenomenon. Building abstractions into material substrates in this way renders the conceptual world available to the senses and makes it possible to manipulate concepts through bodily action. When a skilled practitioner is acting in such a culturally constituted setting, the contents of the generative predictions of the sensory and motor systems are not simply about the location of the parts of the body; they are also about concepts. Thus, a culturally conventional composite abstraction is made accessible to thought in the physical form of the hoey.

The marks on the scale are experienced as directions in space with respect to a reference direction. Culturally conventional resources are mobilized to see the marks AS a particular kind of meaningful object. Their meaning is enacted. The marks become a representation of direction for the navigator in the moment in which they are seen as numbers that denote direction. This culturally conventional composite abstraction (number sequence mapped onto spatial sequence and wrapped into a circle) is pervasive in the navigation activity. It is the foundation for the expression of a direction as a

heading or a bearing. The same composite abstraction underwrites the expression of positions on earth as an intersection of a latitude and a longitude.

Simultaneously, though, the hoey has physical properties that are not direct expressions of the target abstraction. It has a size and a weight, a level of transparency, smoothness to its surfaces, the friction of the base as it slides on the chart, and the friction of the locking ring on its bolt, among others. In a formal sense, these physical properties are implementational details that do not bear on the abstract manipulation of angles. However, because the tool is engaged by an embodied MGC system, these properties enter into the computations performed using the tool in sometimes surprising ways.

Beyond the domain of navigation, the protractor scale is a domesticated space that is linked to thousands of others in our everyday experience. The culturally conventional clockwise direction of the wrapping of the number line suggests an obvious connection. With a different range of values interpreted as times rather than as angles, this same culturally conventional composite abstraction composes the analog clock face. This same abstraction also underwrites a wide range of gauges and dials. We will return later to the cognitive significance of the existence of entire families of tools and practices that share an underlying composite abstraction.

## 5.1.2. Setting the hoey arm to the location of the bearing on the scale

*P loosens the hoey lock with his left hand. His right hand moves toward the hoey base while the left hand rotates counterclockwise to further loosen the hoey lock. Here is smooth simultaneous motion with the two hands engaged in two different aspects of the task.*

*In the last few cm of motion as the right hand arrives at the edge of the protractor, P pushes the hoey arm away from his body (toward lower scale values), while pressing the base into the chart to inhibit its rotation, to align roughly with the scale value 057. There follows a fine adjustment of the position of the hoey arm on the scale value. P overshoots the target by perhaps a few degrees (scale tics are about 2 mm per degree apart), then pulls the hoey arm back to the correct position.*

**Figure 5.1.2. Setting the hoey arm to the bearing on the scale**

*Frame from a video by the author*

This is interpolation in action.

The bearing 057 is not on a labeled tic mark. Interpolation allows a navigator to give labels to unlabeled tic marks. Complex hand-eye coordination (relations among motor and visual modal representations) is required to align the arm with the appropriate value on the scale. The observed overshoot might have taken the hoey arm momentarily to 055 (also not labeled, but a large tic, easy to locate and identify). From 055, pulling the arm up the scale (clockwise or toward the body) by two medium-sized tic marks positions it at 057. This requires counting or perhaps subitizing, which is another complex cultural skill, involving the coordination of a number sequence with discrete shifts of attention.

The continuous interaction of the MGC system with culturally organized material structure permits conceptual work to be done by perceptual and motor processes. In this case, the conceptual work of assessing the difference between the presently set value of the hoey arm on the protractor scale to the value of the bearing to the landmark can be accomplished by visual inspection of the hoey protractor and plotting arm. That difference is reduced by acting on the world, moving the hoey arm across the scale in the direction of the tic mark that denotes the desired bearing. This very simple example points to an absolutely crucial phenomenon in human cognition. Important conceptual distinctions are embodied in salient perceptual distinctions. One can observe, multi-modally, the gradual elimination of the difference between these two quantities.

## 5.1.3 Locking in the Bearing

*Once the arm is positioned on the correct value, P uses his left hand to rotate the locking ring clockwise to tighten the screw and lock in the position of the arm. While tightening the lock, he holds the hoey arm with his right hand to prevent the rotation of the lock ring from moving the hoey arm off the selected bearing.*

The utility of the odd pinching configuration of P's thumb and first finger on the locking ring noted in the previous section now becomes clear. It allows P to impart a clockwise rotation to the locking ring by drawing his thumb toward his hand while extending his index finger by straightening the first joint. No rotation at the wrist is required.

The clockwise rotation to tighten the hoey lock is another cultural convention. The screw in the middle of the hoey base is distantly related to the circular scale around the perimeter of the hoey base. The screw has increasing tension/tightness with clockwise rotation and the scale has increasing numeric values in clockwise rotation. As a consequence of experience with this and thousands of other screw-type devices, the procedure is built into the plotter's body. His body is enculturated. Bourdieu (1977) called these unconscious bodily dispositions and skills *habitus*. The *habitus* of operation of a cultural tool is itself a cultural artifact. This residue of cultural activity accumulates in the body of the experienced practitioner.

Cultural conventions provide predictability, which is just what a predictive system, such as a person needs. Conventions create similar features and similar skill requirements

across task environments, and this supports generalization. The clockwise-is-tighter convention links the plotting activity to those thousands of other activities that involve screw-type devices. The same habits of bodily action that serve the handling of the hoey also support action in a wide range of other activities. Cultural conventions reduce the uncertainty of experience because they increase the uniformity of experience across contexts.

Other properties of the screw lock on the hoey are regular and cultural, but not conventional in the same sense as the clockwise/counterclockwise convention. Consider the slowly increasing resistance of the ring as it is tightened. When the screw is loose almost no torque is required to turn the ring. As it is tightened, additional force is required to turn the ring. Skilled performance requires judging this torque. The screw needs to be tight enough to hold the plotting arm in place with respect to the protractor but should not be subjected to so much torque that the screw's threads are stripped. This is an emergent regularity in the behavior of a cultural object.

The locking ring "locks-in" the value of the angle. Holding the value of a variable constant is a conceptual task. It is accomplished by physical properties of the tool - in this case, increasing friction. Locking in the numeric value of the bearing to the landmark is a form of memory. This is not a memory at the level of the individual plotter's memory, but it is a memory in the wider socio-technical system that implements a physical symbol system.

When used to plot an LOP, the hoey is a digital-to-analog converter, and when locked, the angle is saved for future use in analog form. The importance of converting the concepts to analog form is that the operations on the concepts can be analog rather than digital. Rule-based formal operations on digital representations are a historically recent innovation and are not among the natural strengths of humans, as discussed in the MGC framework section. I will return to this theme later in a discussion of the role of symbols in human cognition.

## 5.1.4 Aligning the Hoey with the Chart's Direction Frame

*P's torso and head begin to move back from the table and up, widening his visual perspective. His left hand opens over the hoey lock and begins to push the hoey base north (away from the body). His left thumb remains resting on the edge of the hoey lock*

*ring, and the fingers of his left hand extend over and rest on the left quadrant of the protractor scale. Simultaneously, P's right-hand fingers spread and anchor the hoey arm, preventing the arm from being pushed away from the body. This results in the rotation of the entire hoey-plus arm assembly around the pivot point formed by P's right-hand fingers. All of this is performed as a single rapid, fluid coordinated motion of both hands and head.*

**Figure 5.1.3. Aligning the hoey with the chart's direction frame**



*Frame from a video by the author*

This is the transition from setting the bearing angle on the hoey to aligning the hoey protractor with the chart direction frame. This rotation produces approximate alignment

of the hoey base with the directional frame of the chart. In this and the following panels, we can see the partial decomposition of the task of getting the hoey aligned simultaneously with the direction frame of the chart and the depiction of the landmark.

*As the hoey base reaches approximate alignment with the chart, P's right hand begins to move out along the hoey arm away from the hoey base. Three actions are happening simultaneously here. 1) Moving the hoey base northward toward the ship's track and the previously marked expected position. 2) Moving the right hand and the hoey arm toward the chart depiction of the landmark. 3) Changing the grip on the pencil from a standby (over thumb, under index, over middle finger) to a conventional writing grip.*

The chart, like the hoey scale, is another domesticated space. With a north-up chart, north is away from the body, south is toward the body, west is to the left, and east is to the right. Meaning in space is experienced by proprioception as well as vision. North-up is another cultural convention (Just ask an Australian), as is the conventional grip for holding a pencil when writing.

In *Cognition in the Wild,* I address the problem of reconciliation of chart to world when traveling in a direction other than north. On this harbor entry, the ship's heading happens to be near north, so motion away from the body is also conceptually forward on the ship's course. This simplifies some computations that are realized in bodily motion. The mapping from body-space to chart-space to ship-space requires only minor mental rotation.

There is an interaction between the physical properties of the motor system and the material world on one hand, and the organization of thought on the other hand. In this case, the ease of some computations that are realized in bodily motion (the mapping between body space and meaningful directions in chart space) is facilitated by a fortuitous correspondence between the ship's present course and the cardinal directions of the chart. It is easy to think "ahead along the course" because that is away from the body. Objects to the left of the ship's course in the world are to the left of the ship's course as enacted by bodily motion on the chart. Objects to the right in the world are to the right on the chart. These simple relationships do not hold when the ship's course is any direction far from true north. In order to reconcile such relationships as enacted on the chart with the situation of the ship in the world requires mental rotation.

There is a more general point here. There are real advantages to material anchors for conceptual entities. However, interactions with material anchors also open the door to

properties of the physical constraints on courses of action affecting conceptual processing. The representation of some situations on the chart may require more or less processing. For example, working with a ship's track on a north-up chart when the track is southward demands more processing resources than when the ship's track is northward. This interaction between motor processes and thinking with material anchors already showed up in the use of the hoey and will also appear later in examples of the acceleration profile of the movement of the plotting tools to a target location.

This vignette contains another fortuitous spatial situation. The plotter is right-handed, and for plotting this landmark, the landmark is to the right of the course. The hoey base can be aligned with the chart right side up (as indicated by the orientation of the text and numeric symbols on the hoey and the conventional orientation of a protractor - 000/360 at the top), and the landmark is to the east (right) of the projected entry track. When plotting a line of position from a landmark that is above and to the left of the estimated position, the navigator must turn the hoey base upside down to align it with the direction frame of the chart and must use the inner semi-circular scale to set the bearing.

*While P's right hand moves out along the length of the hoey arm, his left hand pushes the hoey base away from his body. His torso and head initially follow the base, then his head and gaze shift to his right hand. His left hand stops pushing the hoey base away, and his left hand begins to open three frames (0.1 second) after his gaze leaves the base.*

Here we see the translation of the hoey without rotation to establish a relationship between the depiction of the landmark on the chart and the previously plotted estimated position. The movement of the hoey on the chart surface while seeking an alignment between landmark depictions and the estimated position of the fix raises an important question: To what extent is the location of the hoey on the chart experienced as a location in ship space (having meaning about locations in the world around the ship) and to what extent is it simply experienced as a location on the surface of the chart (in chart-table space) near where the plotting action will happen? This is important because if we are being careful, we should not say that the chart is a representation when it is engaged as a thing-in-itself. The balance between the salience of the two frames of reference must shift as task characteristics change.

This bears on the two aspects of the chart as a representation: The chart as a thing-in-itself versus the chart as a representation of something it is not - the space around the ship. Is it possible to find behavioral indications of which stance the plotter is taking? One clue is this: would we say that the hoey is moved "away from the plotter" (chart space), or would we describe the same motion as "north" or "along the ship's track" (ship space)? Of course, what matters here is how the plotter experiences this, not what we would say as observers. Given the constraints on our video observations, we can recognize these two distinct possible stances with respect to the chart as representation, but we cannot say with confidence whether the plotter is experiencing the chart as a representation or as a thing-in-itself, or as both, in any particular moment.

Once the bearing is locked into the hoey, the plotting task can be undertaken entirely as mechanical procedures on the chart space. The chart will be a representation in some moments and a thing in itself in other moments. And perhaps it is both in some moments. While seeking a landmark that produces a LOP that crosses the track in about the right place, the chart is likely seen as a thing in itself. It is most clearly a representation of space around the ship in the process of choosing landmarks.

## 5.1.5 Evaluating the Relation of the Landmark to the Estimated Position as Implied by the Hoey

*The heel of P's right hand contacts the chart surface at the end of the hoey arm, which is over the depiction of the Dive Tower landmark. P brings the pencil point down to the surface of the chart under close gaze. Body, torso, and left-hand motion have ceased. Impaling the landmark depiction with the pencil point requires fine motor control. Other motion is stopped.*

The heel of the hand provides a steady platform for fine aiming of the pencil point by finger movements. Stopping motion provides a stable frame of reference for visual, motor, and proprioceptive experience. Visual, motor, and proprioceptive expectations can settle into stable states. This cessation of motion **is** concentration.

*P stabs the chart with the pencil point, then lifts the point and moves it north, then comes back near the original point and contacts the chart again, all the while keeping the heel of his hand anchored on the chart. Minor movements of the head and torso shadow these movements of the hand.*

*P's shifts his gaze back to the hoey base in anticipation of the precision motion of the base into alignment with the chart. P then makes small adjustments to the hoey's alignment with the chart directional frame.*

One of the red reference lines in the hoey base must be superimposed exactly on a line of latitude or longitude on the chart while the edge of the hoey arm passes over the depiction of the landmark. There is a choice to be made from a large number of alternatives. There are 13 vertical red lines and 8 horizontal red lines on the hoey base. One of the vertical lines on the hoey base must be superimposed to align with a line of longitude on the chart, **or** one of the horizontal lines must align with a line of latitude. In either case, the alignment must be accomplished in such a way that an unobstructed section of the plotting arm passes over the projected ship's track in the vicinity of the estimated position. How is this choice made? P must get the base into approximate alignment with the chart. He must move the base toward and away from the LM while maintaining contact of the plotting arm with the LM. As this is done, one of the lines will come into alignment with a line of longitude or latitude. This search for alignment is visible on the video. It is subtle because the squares on the hoey base are 13mm on a side. The maximum distance required to align one of the lines is sqrt(2.0)*6.5 = 9.192 mm. Less than a centimeter of displacement is needed to get an alignment. The MGC theory applied to the details of the task requirements and the physical properties of the tools predicts this subtle movement. Once one knows to look for it, it is visible in the video. Without this prediction, the tiny motions would probably be disregarded as noise, and the insights would elude the analyst.

While adjusting the alignment of the base with the chart, the hoey arm must be kept in contact with the pencil point. This is accomplished by keeping gentle counterclockwise rotational torque on the hoey base as it is moved around. Downward pressure on the hoey base must be relaxed in order to allow the hoey base to slide easily and smoothly over the surface of the chart. The skills of an expert navigator are not just a matter of knowledge; they include very finely tuned motor routines.

## 5.1.6. Detecting an Error

*There is a moment of virtually no body motion while P looks at the hoey base. P's eyes are not visible, but I presume that they move from the reference lines in the hoey base to the edge of the hoey arm where it crosses the projected ship's track. This is where*

*the LOP should be plotted. It happens to be very close to the edge of the hoey base, so no change in head orientation is needed to see it.*

**Figure 5.1.4. Detecting the error**



*Frame from a video by the author*

The hoey arm crosses the track line south of several previously plotted positions. By spatial inference, this cannot be the correct position for the LOP. The navigator can imagine the simple dynamics of the ship moving forward along its track northward from the last position fix. The correct position must therefore be much further north, up the track line (which bears about 10 degrees W of N). What is the nature of this spatial inference? It must include imagination of the progress of the ship along its track. It

requires the superposition of a generated trajectory on a perceived spatial array. This happens at the meeting of generated structure with sensed structure. Neither the generated trajector of attention alone nor the perceived ship's track alone is sufficient to make this inference. The terms of the inference are emergent from the ongoing projection of the trajector onto the array as it is being perceived. This is the same cognitive routine that produces number lines and the sense (root meaning) of the direction of increasing values on the protractor.

This type of error could be the result of plotting the wrong bearing or projecting the correct bearing from the wrong landmark. The plotter does not seem to explore the possibility that the bearing is wrong. We cannot know for sure why this is. The correct landmark for the 057 bearing is Hotel Del Coronado. This is how it was originally reported by the BTR. Recall, however, that when the bearing was reported with the landmark name, P was distracted by a question from the keeper of the deck log. When he asked for the bearing again, BR told him the bearing, but did not repeat the name of the landmark. Possibly because of the distraction caused by the question from the keeper of the deck log, the plotter first tried to fit the LOP to a landmark called Dive Tower. The LOP from this landmark did not work. Patterns of information flow among the watch-standers on the ship's bridge affect both the cognitive processes of individuals (in this case, vulnerability to error) as well as the cognitive properties of the bridge team as a cognitive system.

*A moment without motion marks the discovery of the error.*

The words we use to describe the event, "discovering the error," are not the thought. The thought that the LOP must be plotted with respect to a different landmark north of Dive Tower is not a mental proposition. The thought that there is something wrong with the LOP is the entire dynamic multimodal experience.

It includes:

- The visual process is a continuous engagement with the chart and the hands, and the hoey.

- The tactile experience of the roughness of the knurled knob, the smoothness of the chart and the hoey surface.

- The proprioceptive sense of the location of the hands in body-space, in chart-space, and perhaps in ship-space.

- The sense prediction that pushing the hoey away (north) will bring the hoey arm to the EP.

The thought is this entire continuous dynamic experience of the body and the world. There is intimate contact of the body and nervous system with a thoroughly culturally organized world, adapting, resonating, and entraining. Coupling prediction to sensation to perception, and modality to modality. Parts of this thought are in the world, and some of those parts are even visible to us as observers.

## 5.1.7. Correcting the Error

*P's head begins to come up. His right hand lifts the pencil off the chart surface.*

*This is the beginning of the abandonment of the previous (mistaken) solution. It does not appear that the plotter's gaze goes up-track toward the EP for the current fix. Rather, it goes toward the plotter's right hand, which is already in motion up the chart away from his body (northward). Searching for a workable landmark.*

It is not possible to know if the plotter noticed the approximate distance on the chart from the incorrect LOP location (implied by Dive Tower) to the previously plotted EP. It seems likely that he did, because he certainly evaluated the incorrect LOP location and noticed its directional relation to the previously plotted expected position. If he did notice that distance, he could be looking for a landmark that is north of Dive Tower by about the same distance. He does seem to move quickly to Hotel Del, which satisfies that constraint.

*P's head snaps toward his right hand, and his left hand begins to push the hoey (base and arm with almost no rotation) north, away from his body.*

**Figure 5.1.5. Correcting the error**

*Frame from a video by the author*

This is the initiation of the corrective action, searching for a workable landmark. P's head began to rise just two video frames (1/15th of a second) before his hands began to slide the hoey base away from his body. He appears not to have questioned the bearing or the setting of the hoey arm locked into the hoey base.

*The pushing motion accelerates. P's head reorientation ends. His right hand rides along on the end of the hoey arm, keeping pace with the left hand, which is sliding the hoey base. Thus, the hoey maintains its approximate rotational orientation with the chart. The pencil remains upright in P's right hand during this translation, ready to impale the new landmark. The motion of the hands decelerates as the hoey arm and right hand arrive in the vicinity of the depiction of Hotel del Coronado. P's gaze is now fixed on the*

*chart under the pencil point. His torso continues to advance over the chart, catching up with his hands and bringing his head closer to his right hand.*

P searches the space of possibilities by moving the entire hoey away from his body (north) without rotation until the segment of the plotting arm closest to the hoey base is near the estimated position. There must be a familiar proprioceptive signature of equal displacement of the two hands to produce movement without rotation. This is search and choice carried out in actions in a domesticated space.

Here, spatial reasoning is enacted in the coordination of visual, spatial, and motor modalities. It is not known whether P thought of Hotel del Coronado at the moment of discovery of the error, or if this landmark was noticed later when the corrective motion was begun. Since his head motion ends near the depiction of that landmark, it is likely that P has the depiction of Hotel del Coronado in sight by this time. The orientation of P's hoey with respect to the chart is maintained in this lateral translation.

The entire solution is roughly described by this single motion of the hoey up the chart. While the motion of the base does not seem to be the controlling parameter in this motion, P has somatosensory proprioceptive indications that the motion of the base is in the correct direction and, keeping pace with the right hand, it will maintain the rotational orientation to the frame of reference of the chart.

The velocity profile of the movement of the hoey is as expected for a movement stroke in motor control theory. There are two interpretations for the velocity profile. First, this might simply be a product of the physics of the motor system, smooth acceleration followed by smooth deceleration (Flash & Berthoz, 2021). This is a relatively simple interpretation in which the context provides only starting and ending points. Second, the motion accelerates where vision indicates there are no suitable landmarks and decelerates when approaching the depiction of Hotel del Coronado. This is a more complex interpretation in which velocity is dynamically and continually controlled in a feedback loop that involves interactions between the visual modality and the motor system. On the basis of the available data, it is not possible to determine whether one or the other of these interpretations of the velocity profile is dominant. It is possible that both are involved in shaping the velocity profile.

In any case, elements of this thought are visible and measurable actions in the world. This raises the question of whether the physical properties of the motor system might affect thought processes where courses of action are trains of thought. To the extent that

we think with our hands, our thinking may be constrained by the physical limitations on hand motion. Resolving the possible effects of motor system dynamics on the dynamics of thought will require other methods.

Bringing his head closer to his right hand facilitates the scrutiny of the pencil point and the landmark depiction. The pencil is held upright - bodily state anticipates the next action (bringing the pencil point to the chart surface), which will be the next concept (location of the landmark depiction).

*For a beat, P holds still, presumably studying the chart.*

The depiction of the Hotel del Coronado complex has some detail, and the labeled tower must be located among other buildings.

## 5.1.8. Verifying the Landmark Identity

*P's head begins to turn to the right, initiating a glance at the bearing record log. His right arm begins to straighten and stiffen as his gaze arrives at the bearing record log. He keeps the pencil point in contact with the chart at the Hotel del Coronado landmark.*

Before locating the Hotel Del tower with precision, the plotter checks the identity of the landmark. It is clear that he did not suspect that the bearing was incorrect. His first corrective action was to move the hoey to a different landmark, and the second was to check the identity of the landmark in the bearing record log. There is no indication that he looks toward the Estimated Position to see if this LOP will make sense with respect to the EP. Straightening the right arm as he twists his torso rightward, following the motion of his head, allows P to keep his right hand in place near the Hotel Del landmark depiction. At the system level, the current position of the hoey on the chart is a memory for the current state of the attempt to solve the problem. Some rearrangement of the body is required to maintain that state while twisting to read the relevant entry in the bearing record log. This embodied memory for the current state of the problem solution across an interruption.

Consulting the bearing record log allows the plotter to find the bearing 057 and read the landmark name at the top of the column in which it occurs. Visual search is a complex and well-studied problem. It is beyond the scope of my work. The plotter scans the

bearing log with a visual expectation that is shaped by an earlier auditory input - the spoken bearing - and the visual and motor experience of having set the bearing on the hoey protractor. Once the written bearing has been located in the bearing record log, P needs to scan up the page to find and read the name of the landmark at the top of the column.

The row and column format of the bearing record log is a cultural device that controls the relationship between two variables. This cultural invention first appeared about 4,000 years ago (Marchese, 2011). Its physical organization affords embodied interaction. The procedure could be described in three steps. Step 1: Scan down the time column to the desired time (again, that projected trajector). The current time is easy to find because it is the last one written in the time column. The cell below it is empty. The conceptual work of finding the most recent time is a simple embodied action. This action could treat the time column as a representation of the passage of time, or not. The remaining actions treat the table as a thing in itself, not as a representation. Step 2: Scan across the columns to find the bearing 057. Step 3: Scan up to the top of the column (backward in time, but time is not relevant to this action) to find the landmark name.

## 5.1.9 Drawing the LOP Segment on the Predicted Track

*P turns his head to the left and his gaze returns to the hoey base. There is little other motion except subtle adjustment of his torso. He holds the hoey in place with the fingers of his left hand on the surface of the protractor and the heel of his hand firmly on the chart. P moves the hoey base slightly toward his body, aligning it with the chart. The hoey arm remains in contact with the pencil at the Hotel Del Coronado landmark.*

*P slightly relaxes and contracts his left hand on the surface of the hoey. Then he is almost motionless for two seconds. He wiggles the pencil on the landmark and makes very subtle motions of the hoey base aligning with the chart. P moves his right hand to the intersection of the hoey arm with the previously plotted anticipated ship's track. His head rises slightly, but his gaze never leaves the target area for drawing the LOP. P brings the pencil into contact with the chart to the left of the ship's track. P draws the LOP as a short, sharp stroke across the ship's projected track.*

**Figure 5.1.6. Drawing the LOP**

*Frame from a video by the author*

Without eye-tracking recordings, it's impossible to know exactly what P is focusing his vision on. The estimated position is in his field of view just to the right of the hoey base. The wiggling of the pencil and the subtle translations of the hoey base are fine-tuning the position of the hoey. Conceptually, this amounts to refining the constraints that will be embodied in the plotted line of position.

The line of position could be drawn the full length of the hoey arm. All of the positions on such a line fall on the bearing from the ship to the landmark. However, P has strongly constrained expectations about where the position fix will fall. He only needs to make this line long enough that it will intersect the lines of position for the other two landmarks. Any length beyond that simply clutters the chart. So, the LOP is drawn as a short line near the anticipated ship's track.

Notice the acceleration of P's hand in drawing the LOP. If the thought is the action of drawing the LOP, and I believe it is, then again, the dynamics of bodily motion may embody and constrain features of the thought. The LOP is drawn with a flourish characterized by rapid, smooth acceleration culminating in a sharp stop. It seems to me, as an observer, that the smooth acceleration represents the course of the procedure, and the sharp stop represents its completion. I suspect that some other investigative techniques could resolve such questions.

The motion of the body is a visible aspect of the thought. What is the relationship between the dynamics of the motion and the dynamics of the thought? The same question was raised by the acceleration of the motion of the hoey up the chart, moving from the landmark Dive Tower to the Hotel Del Coronado. To what extent is the velocity profile of such motions constrained by the features of the problem, and to what extent is it constrained by the physics of the body? Bodily actions are more than simply imperfect expressions of an underlying thought that are perhaps corrupted by the expressive limitations of the body. Motion and thought have a reciprocal influence when thinking is done with the hands.

P stops all other motion in the body when he focuses on fine work with the pencil. This seems to be the physical manifestation of concentration and focus of attention. Perhaps the dynamics of action ARE the dynamics of thought. If this is true, then the properties of the motor system, as a physical system, may constrain thought.

A moment without motion marks the discovery of the error. When courses of action are trains of thought, interruptions in the train of thought may appear as interruptions in the course of action. If we take embodiment seriously, then the space of possible thoughts must be constrained by the physical characteristics of the body.

## 5.2. Using the Three-Minute Rule to Compute Ship's Speed

[7] As we have seen, what have traditionally been called high-level cognitive processes can be produced by the engagement of a culturally orchestrated MGC system with cultural materials, that is, elements of language, sign systems, and inscriptions of all sorts.

A rich 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 a 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 general procedure for computing speed is to measure the distance traveled by the ship over a given span of time and divide the distance by the time. There are many unit systems for the expression of distance, time, and speed. There are also many ways to perform this computation, each one making use of a different configuration of internal and external cognitive resources. When the distance is expressed in nautical miles and the time in hours, the result is nautical miles per hour, also known as knots.

The three-minute rule depends on a serendipitous interaction between two systems of distance units (the nautical mile and the yard) and two systems of time units (the minute and the hour). The nautical mile is centuries old and comes from scientific efforts to measure the Earth. The coordinate system of the Earth draws on the Mesopotamian base-60 arithmetic of four millennia ago. The circumference of the Earth is defined as 360 degrees of arc. Each degree contains 60 minutes. Thus, there are 21,600 minutes of arc on the circumference of the Earth. A nautical mile is one minute of arc or 1/21,600th of the circumference of the Earth. The yard has a murkier history but seems to have origins in the Saxon culture of about a millennium ago. One legend holds that it was originally a measure of the single-arm span of a particular English king. The hour comes from the Roman tradition of military watch standing two millennia ago, while the minute of time, like the minute of arc, goes back to the base-60 arithmetic of Mesopotamia and divides its containing unit, the hour, into 60 equal parts.

Minutes of arc and minutes of time thus share historical roots. Yards and hours, however, were developed independently of the others over the course of histories that span thousands of years. Entirely coincidentally, a nautical mile is very nearly 2000 yards. By definition, 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. Virtually all ship navigators know this rule and can use it, but few know why it works. The three-minute rule is such a powerful regularity that one doesn't have to know why it works to discover it. Repeated

experience of the sort, "When 1500 yards are covered in 3 minutes the speed is 15 knots", "When 1200 yards are covered in 3 minutes the speed is 12 knots" and so on, compel the induction of the rule "When X hundred yards are covered in 3 minutes, the speed is X knots."

This convenient fact is put into practice in navigation in the following way. Two successive positions of a ship are plotted at 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, pp. 151-152). This is a dramatic compression of computational procedure that is made possible by the opportunistic arrangement of cultural resources. This sort of opportunistic combination of disparate elements is typical of culturally elaborated activities.

In the terms I was using in *Cognition in the Wild*, this high-level cognitive function is realized in the transformation and propagation of representational states. The span between the fix positions on the chart is a representational state that is propagated into a span on the dividers. This analog representational state is then propagated into a span on the yard scale. Finally, the span on the yard scale is transformed into a digital 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 activity through the lens of MGC. This will show that an MGC analysis of the three-minute rule creates explanatory possibilities that simply have no place in a disembodied analysis.

**Figure 5.2.1. Spanning the distance on the chart**

*Photo by the author*

The navigator's first step is to see and apply the dividers to the span of space between the position fixes (figure 5.2.1). 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 a finely tuned habitus of action and perception. This includes 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, locating the next fix triangle visually, and then sticking the second fix triangle with the other point of the 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 5.2.2). The navigator must now raise the dividers off the chart 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. Notice that the two tasks, adjusting the span followed by maintaining the span while moving it, put conflicting demands on the tool. The span 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. We saw a friction lock on the hoey in the example of plotting a line of position.
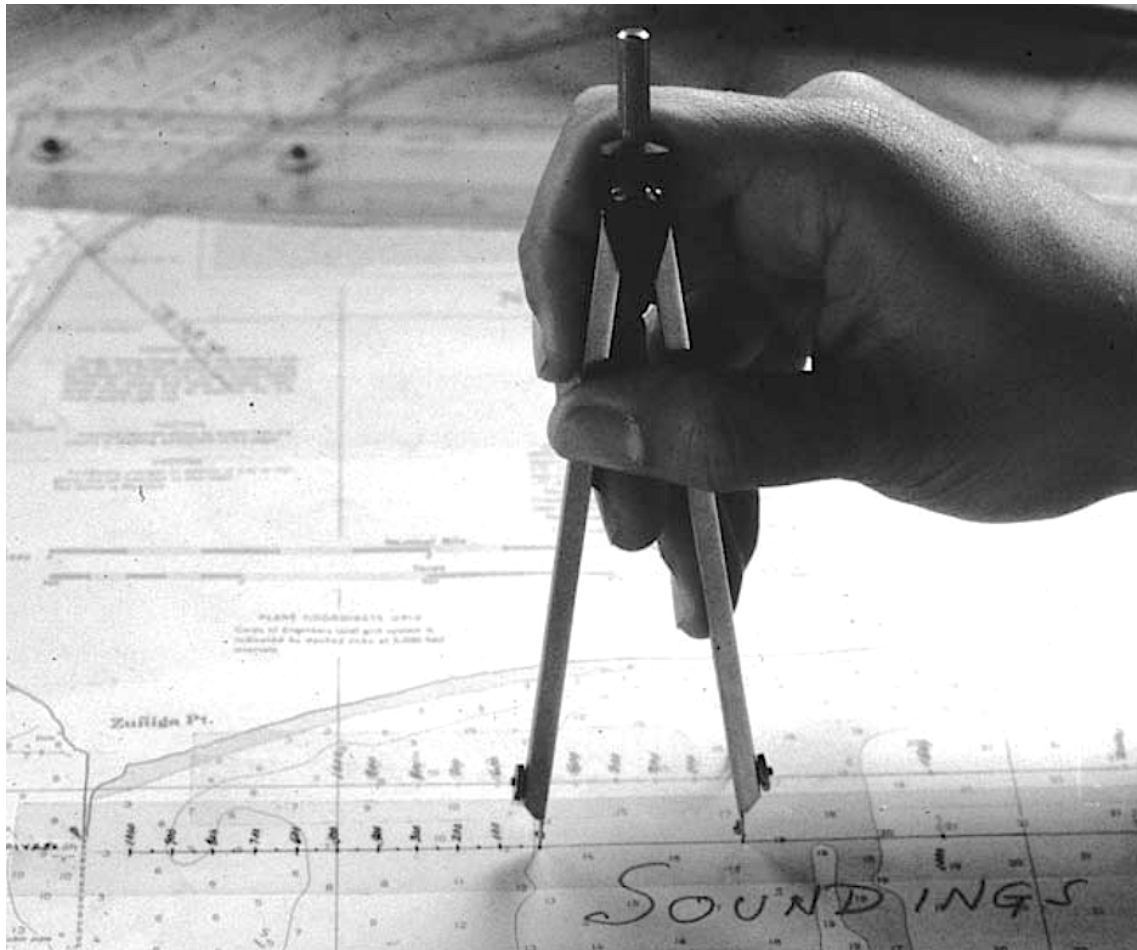
**Figure 5.2.2. Spanning the scale**



*Photo by the author*

Once the divider is placed on the distance scale, the navigator uses the point of the divider arm to direct his attention to the region of the scale under the point. Through this perceptual practice, the divider point is used to highlight (Goodwin, 1994) a position on a distance scale. If the highlighted position is not precisely on a labeled point on the scale, the navigator will have to establish the value through interpolation (as described in the discussion of reading the degree scale on the perimeter of the hoey in the previous section). 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. In this case, by generating a predicted visual sensation that plays the role of distance in

the underlying conceptual network of the predictive processing system. 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 modalities of experience lends stability to the enactment of the measured distance.

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. 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 MGC 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 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. Distance is the meaning the navigator expects when the dividers are in contact with the chart; speed is the meaning the navigator expects when the dividers are placed in contact with the scale. 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 an entire MGC system 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.

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 a particular set of dividers. 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 (visual, tactile, and proprioceptive modalities) and the position of his body (proprioception) 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 multi-modal but is also multi-temporal 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. 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. Sensorimotor contingencies are 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 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 et al., 2003)

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 a culturally organized world.

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, p. 135).

Combining the basic embodiment premise that low-level action and perception are inextricably linked, with the idea from Havelange et al. (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.

# 5.3 An Aha! Insight

[8] As described above, when plotting a position fix, 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 constraints from which the fix triangle is constructed. 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 the Palau, video-recording navigation activities, when, while entering the narrow navigation channel at the entrance to San Diego Bay, 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 of the landmark by adding the corrected magnetic ship's heading to the relative bearing of the landmark (bearing of the landmark with respect to the 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 by the navigation team before sailing, 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. Variation is indicated in the compass roses printed on the navigation chart. The correct equation is as follows: 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 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. Figure 5.3.1 shows our reconstruction of the position fixes created by the navigation team before and after the failure of the gyrocompass.

**Figure 5.3.1. Reconstruction of the position fixes created by the navigation team for the entry on which the gyrocompass failed**
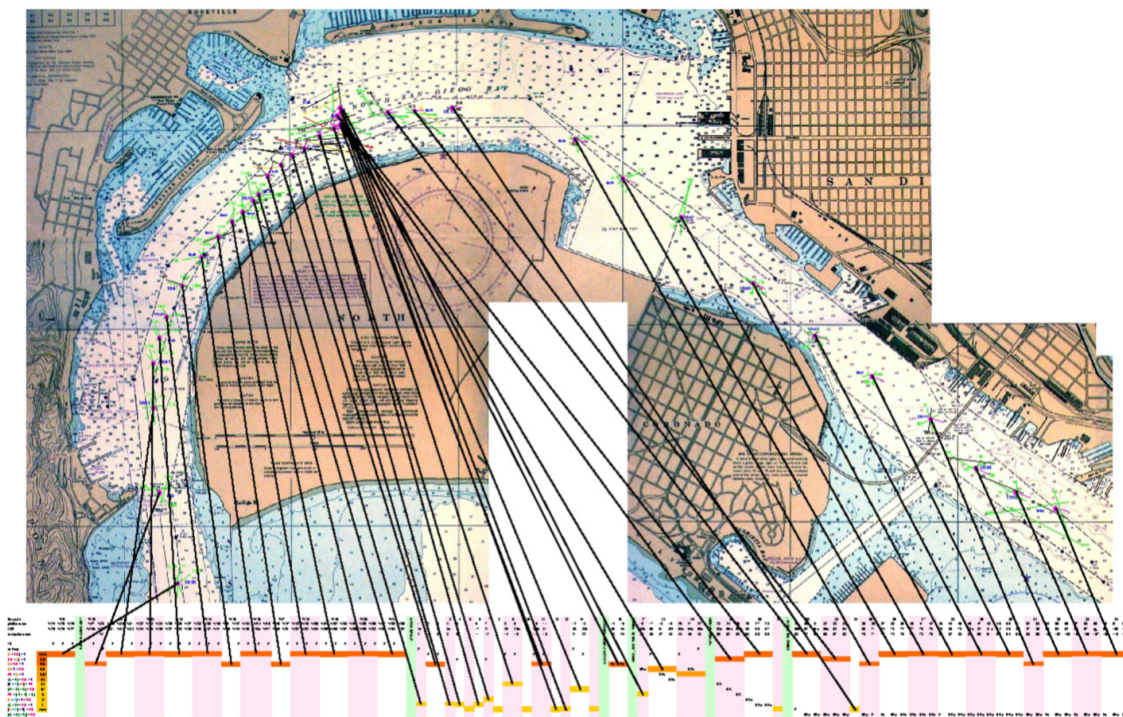
*The column at left indicates the computational sequence used to compute each line of position. The colored horizontal bars across the bottom of the figure indicate the quality of the resulting fix. The green vertical bars indicate key adaptive events in the sequence. The gyrocompass failed at the first green vertical bar. The missing deviation term was discovered at the third green bar.*

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 mono-modal (or even a-modal) language. In Hutchins (1995), I provide a disembodied analysis of this event that fails to explain how the discovery of the missing term was made. In fact, I struggled for many years to develop a satisfactory account of how this insight came about. Things became clearer when I added embodiment to my theoretical toolbox. A careful examination of the way a navigator used his body to engage the tools in the setting 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.

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 P to BR. He said, "I keep getting these monstrous frigging goddamned 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 P.

Such a large triangle was clear evidence of the presence of an error somewhere in the process that created the fix. P checked the addition of terms for the LOPs, 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. P then used the plotting tools and the chart to explore changes to LOPs that might improve the position fix.

Table 5.3.1 has two columns. In the left column are descriptions of the observed 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 a dynamic, multimodal, continuous, 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 P's course of action. 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 5.3.1. P's activity and the enactment of phenomenal objects**

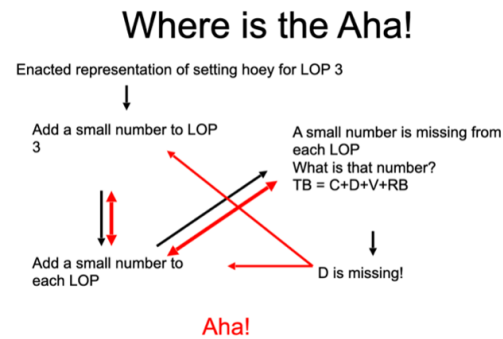| SPEECH | GESTURE IN SPACE |
|--------|------------------|
| but | Pointed with index finger to the engine three fuel gauge |

| Observed activity | Enactment of phenomenal objects |
|---|---|
| P 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 5.3.2, left). | In these first two moves, P 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. |

| Observed activity | Enactment of phenomenal objects |
|---|---|
| P 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. |
| P 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 P 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 the first two LOPs. |

| Observed activity | Enactment of phenomenal objects |
|---|---|
| P 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 5.3.3). 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. |

| Observed activity | Enactment of phenomenal objects |
|---|---|
| P 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: | This action sequence contains more self-regulatory speech. Fidgeting is not noise. Three new concepts have been integrated in this moment. They correspond, in order to the three observed action elements. They are as follows: |
| P 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. | One: that the small number that would improve the LOP is deviation. The deviation table is posted at the helm station. |
| P then came back to the chart table, saying, "say three, say three (accompanied these words with beat gestures), add three to everything." | Two: that all three LOPs will be improved by adding to them a small number. He says, "add three to everything." |
| Upon hearing P say this, BR asked, "Add three? Because we're shooting relative?"<br><br>P responded, "Um, no. On a southwest heading <u>add</u> three." | Three: 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 present heading to the need to add 3 degrees to the LOPs. |
| P then re-plotted the three LOPs, adding 3 degrees to each. This produced the desired small fix triangle. | These three concepts form a synergistic cognitive ecosystem in which each of them makes the others stronger. (figure 5.3.2, right) |

**Figure 5.3.2. P manipulating the hoey (left)**

## Where is the Aha!

Enacted representation of setting hoey for LOP 3

Add a small number to LOP 3

A small number is missing from each LOP
What is that number?
TB = C+D+V+RB

Add a small number to each LOP

D is missing!

Aha!

*Mutual elaboration of concepts on (right). Follow black arrows through a sequence of insights. Red arrows are mutual support*

The tentative nature of these actions 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 an 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.

**Figure 5.3.3. The superimposition of imagined clockwise rotation (motor anticipation) onto the visual experience of the hoey degree scale**
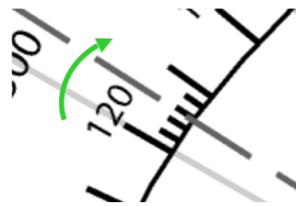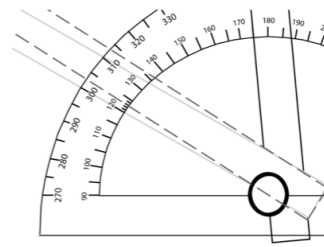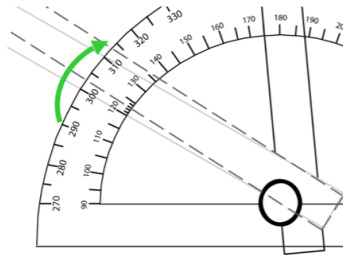
*Light-gray solid lines represent the position of the hoey arm when aligned with the 120-degree mark. Dashed lines represent the imagined location of the hoey arm if it were rotated slightly clockwise. The image of a number slightly larger than 120 is an emergent property of this interaction between the 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 P 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 P are multimodal, generative, and that the contents of the various modalities 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 et al., 1994; Noë, 2004), and the temporal dynamics of many sorts of action are characterized by both feedforward and feedback effects (Spivey, 2007). The match between expectation and sensation is, of course, the key generative element of predictive processing systems. Furthermore, 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, pp. 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, p. 268).

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 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 (Havelange et al., 2003; Hutchins, 2005) 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 remains. 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 here.

The "Aha!" insight here is that the deviation term is missing from the computation of true bearing of the landmark. The MGC 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 offered 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 the clockwise rotation of the LOPs. The visual experience of the protractor scale is a necessary component of the activity the navigator is engaged in. 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 sensation 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 clockwise, that is, to a higher value, on the scale. I believe that the MGC 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.

The processes described thus far can be characterized in terms of some general implications of the MGC view of cognition.

In certain culturally constructed settings, bodily motion acquires meaning by virtue of its relation to the spatial structure of things. Goodwin (1994) 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 a 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 an 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 these reasoning skills as "thinking like a compass." I believe they 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, p. 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 add stability 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. This variability in performance is often considered to be noise that is 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. Variability is important for the dynamics of cultural evolution and for the control of entropy or uncertainty. Little is known about this aspect of cultural systems.

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. Seen through

the lenses of the MGC approach, 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.

Multimodality is a fundamental property of lived experience, and the relations among the contents of various modalities appear to have cognitive consequences. Congruence among the contents of modalities 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 above. 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 MGC perspective reminds us once again 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 enactment 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 can result when culturally orchestrated MGC 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 enactment of integrated, multimodal, temporally extended, embodied representations.

## 5.4 Choosing Landmarks

[9] Having plotted the ship's position and computed its speed, P projected an estimated position (EP) - the anticipated location of the ship at the time of the next fix (shown as a half circle on the anticipated track line in Figure 5.4.1). P and BR now consider landmarks to use for plotting the next fix.

**Figure 5.4.1 Estimated position is the dot in the half circle on the planned entry line**
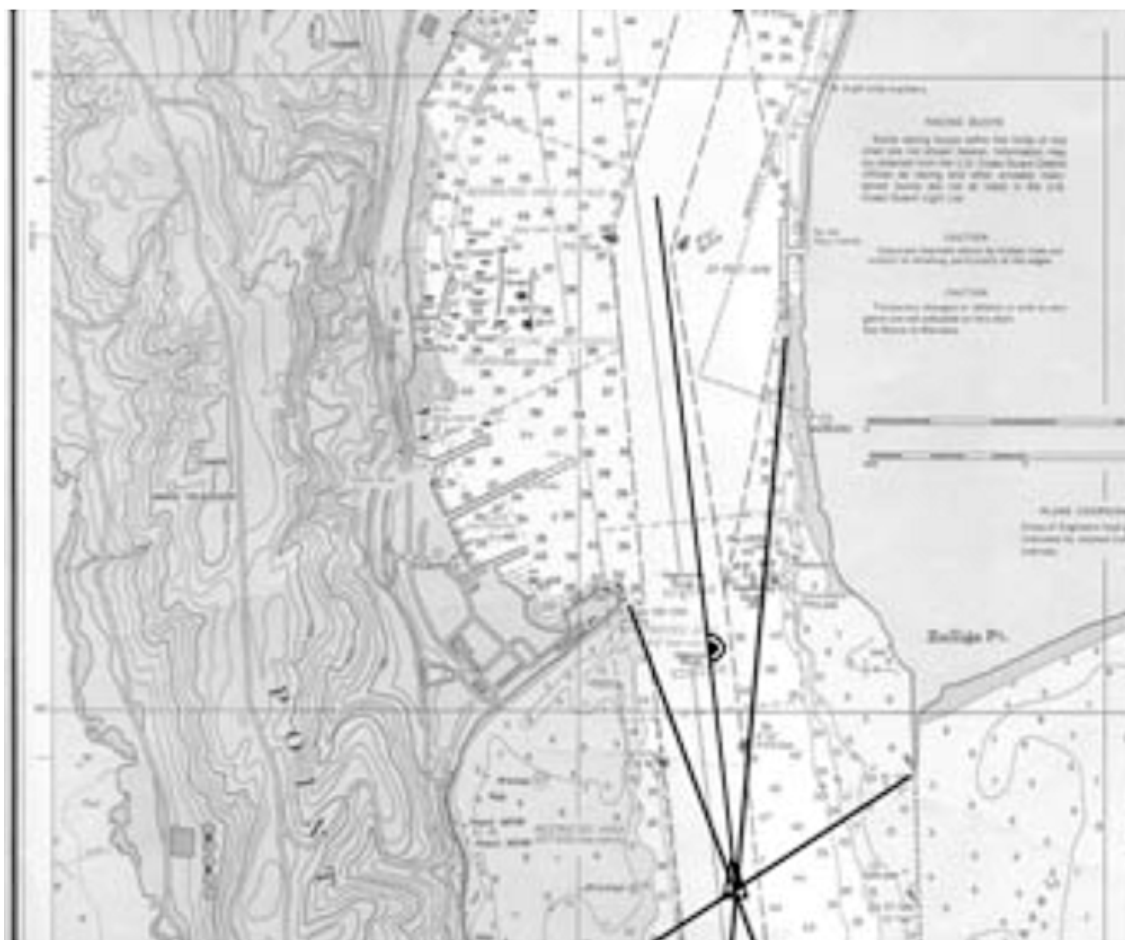
The trick here is that the three landmarks must be chosen such that the LOPs intersect at useful angles (Figure 5.4.2). The choice criteria are as follows: Shallow angles between LOPs are to be avoided because with a shallow angle of intersection, a small error in either LOP will cause a large displacement of the point of intersection between the lines (See the two dashed lines in Figure 5.4.2). Ninety-degree angles are least vulnerable to displacement of the point of intersection with errors in the bearings, but of course, it is not possible to place three lines with ninety-degree intersections in a two-dimensional space. Experienced navigators judge the shape of the triangle, rather than evaluating each of the three intersections separately. A 90° intersection with two 45° intersections is a good solution, as is an equiangular triangle with three 60° vertices.

**Figure 5.4.2. LOP angles**
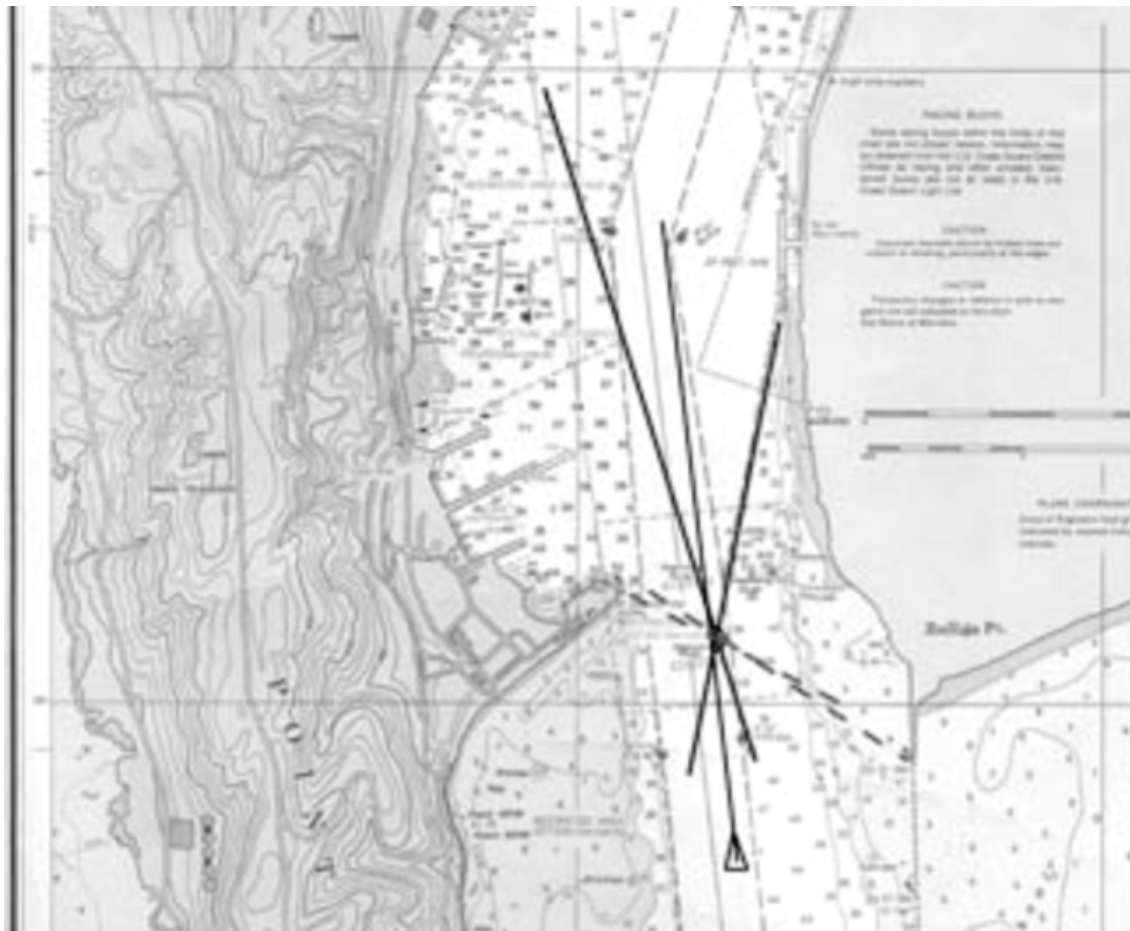
Thus, although a position fix consists of three elements (LOPs), none of the individual elements can be said to be good or bad with respect to the choice criteria. The criteria refer to the relations among elements, not the elements themselves.

This choice process can be taken as a model of a more profound phenomenon to be encountered below. The meanings of elements of multimodal interactions are not properties of the elements themselves but are emergent properties of the system of relations among the elements.

To avoid cluttering the chart and making extra work of having to erase rejected LOP proposals, the navigators imagine the LOPs, rather than drawing them. Once a set of three landmarks has been chosen, the landmarks will be communicated to the bearing takers on the wings. The LOPs will be drawn on the chart only after the bearings have been measured by the bearing takers and reported to and recorded by the bearing recorder. But this implies that the imagined LOPs must somehow be maintained while the relations among them are evaluated. The navigators create imagined LOPs one at a time by gesturing. This has two effects. First, it is an efficient way for each navigator to

communicate complex spatial relations to the other navigator. Second, gesturing seems to create imagined LOPs with sufficient stability to support reasoning about their angles of intersection. The first one created remains available for comparison while the second and third are created.

A transcript of the verbal part of the interaction between plotter (P) and bearing recorder (BR) looks like this:

BR:         so: it'll be that (1.9) n that (1)

P:           Ballast Point (.7) Bravo (1)

BR:         u:[h

P:           [that's good (.5)

BR:         okay (1.2)

The verbal exchange is just one element of the interaction. The interaction is complex in that it involves sensory and motor modalities of both interactants. There is much more to be observed in other modalities. Let's look more closely at the interaction.

*The bearing recorder first proposes two landmarks. He leans over the chart (saying, "It'll be ...") and uses his left index finger to quickly trace a line from a landmark called Ballast Point to the approximate location of the EP (saying "that"). His finger wavers for a moment, making a loose clockwise loop over the chart, then he traces a line from the landmark called Bravo Wharf (saying " 'n that"). BR's left hand remains in the vicinity of the*

**Figure 5.4.3 Choosing LMs**

*Frame from a video by the author*

*P interrupts BR's activity by moving his right hand, middle finger extended, into the area over the chart where the estimated position has been plotted. BR withdraws his left hand from the area as P's right hand comes in. Quickly tracing the imagined LOPs from each landmark as each is named, P revisits the same landmarks just mentioned by BR, "Ballast Point, Bravo."*

*BR tries to retake the floor, leaning over the chart and reaching toward the plotting area with his left hand, saying, "u:h,"*

*P rebuffs BR, making another gestured LOP from the vicinity of the depiction of Light Victor to the EP and saying, "that's good."*

*Because Light Victor is located to the east (right) of the EP, this gesture both indicates a third LOP and effectively blocks the entry of BR's hand to the plotting area. BR pulls his left hand back, rests it on the chart table in front of him, and says, "Okay."*

As the navigators work, they use their fingers to trace lines from various landmarks to the vicinity of the EP. The ephemeral gestures together with the generative apparatus that produces them ARE the provisional LOPs. The provisional LOP is both produced and perceived in a unitary system that includes the chart, the gesturing body, and the internal MGC system.

The ephemeral gestures enact imaginary or provisional LOPs. These ephemeral structures are the representations on which the choice process operates. The creation and evaluation of the proposed LOPs is carried out in a conversation between BR and P. The conversational turns are multimodal in that they include environmentally coupled gesture, co-gesture speech, body orientation, facial expression, and tool manipulation.

The gesture is complex. The hands of the participants move around a lot over the chart (Figure 5.4.4). Some parts of the gesture stream are meaningful. Some are not. Some gestural strokes represent lines of position, whereas other strokes reposition the hand to begin a meaningful stroke. How do the participants distinguish the meaningful parts of the gesture from the parts they should disregard?

**Figure 5.4.4 Gestural trajectory**
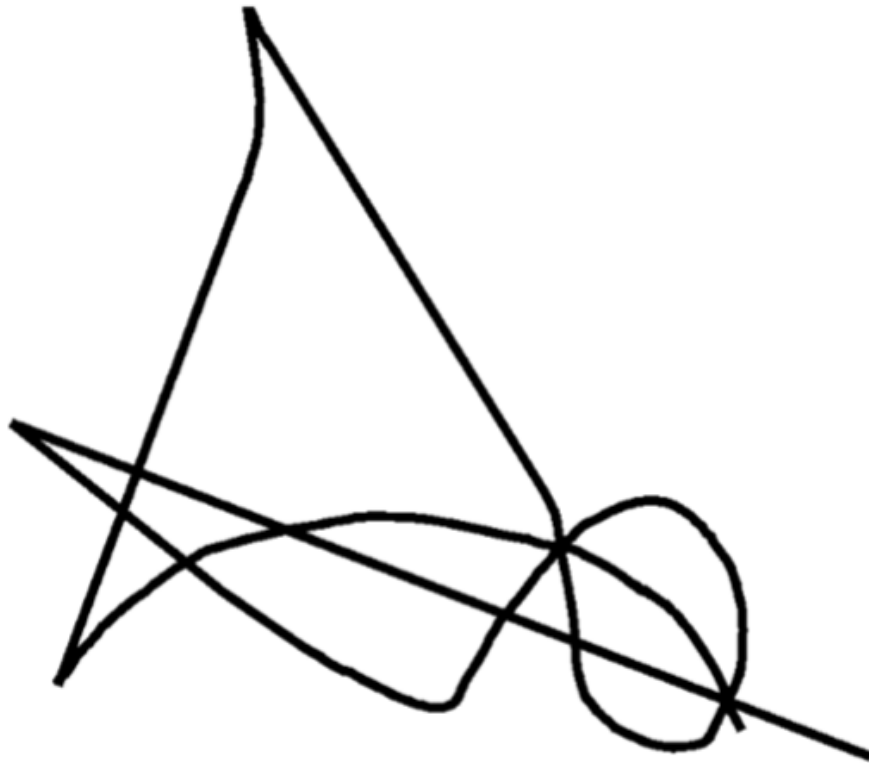
Multiple cues contribute to the identification of parts of the gesture that delineate LOPs. The participants know that the objects of interest are virtual lines of position. These lines should link landmarks with the ship's estimated position. BR says, "It'll be that 'n that." The seemingly unbound anaphora of "It" refers to the object of the understood current project that is the creation of a triplet of landmarks to be used in plotting the next position.

The trajectory of BR's gesture suggests some lines that pass near landmarks and the EP, but it does not unambiguously pick out the potential lines of position that are being proposed.

BR's gesture has a velocity profile. BR's hands move fast in some parts of the trajectory, and they move slowly in other parts. In others, the hands come nearly to a stop. Meaningful gestures often come in the form of strokes that are demarcated by pauses before and after the meaningful stroke. These are called pre- and post-stroke holds (McNeill, 1992; Streeck, 1993). Figure 5.4.5 shows a tic mark at the location of BR's fingertip in each frame of the video. Sparse tic marks indicate rapid motion,

whereas dense areas of tic marks indicate slow motion. This velocity profile suggests pre- and post-stroke holds for two gestural strokes: one on the east-southeast-ward stroke from Ballast Point through the EP, and the other on the south-southwest-ward stroke from Bravo Wharf through the EP.

**Figure 5.4.5. Velocity profile of the gesturing finger**
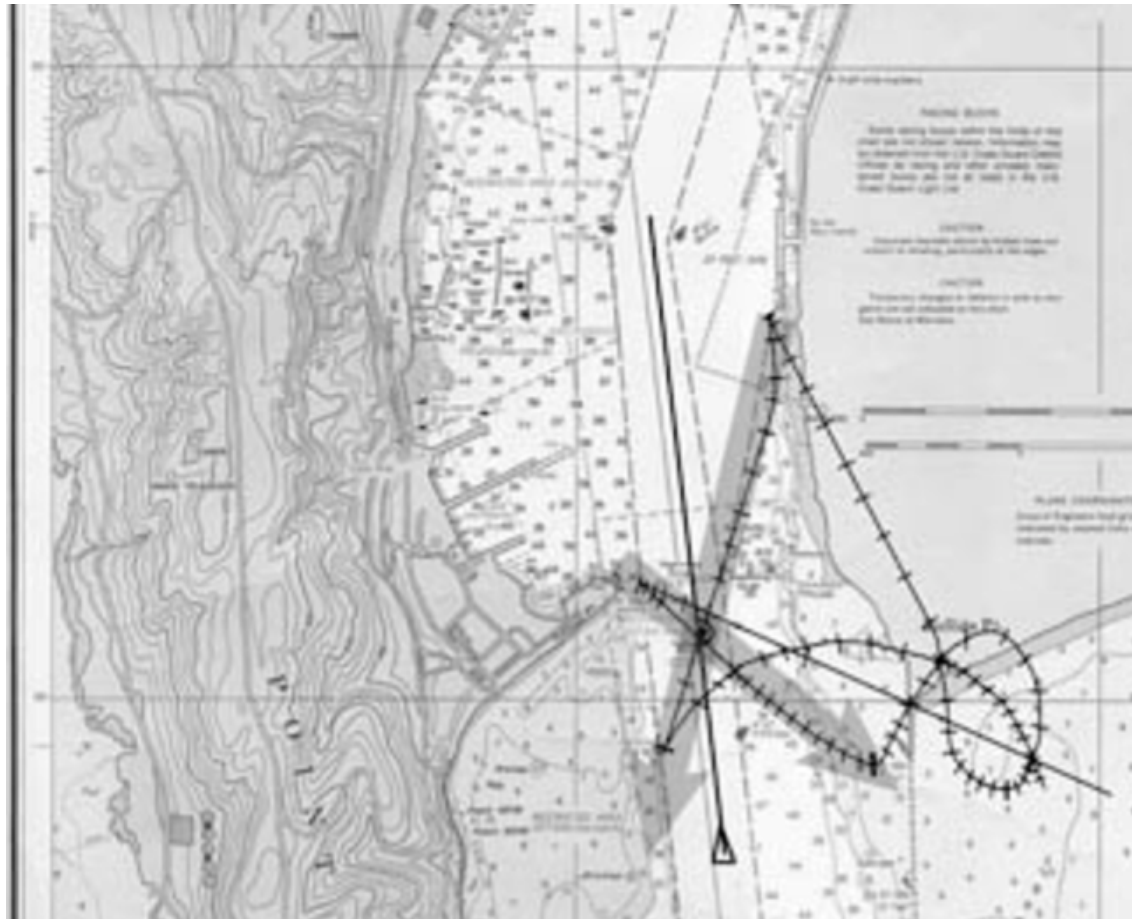


Another useful cue is the shape of the gestural trajectory. Because LOPs are, by definition, straight lines, gesture segments that are curved are unlikely to be meaningful depictions of virtual lines of position. This cue is not, by itself, sufficient to pick out the proposed LOP's. In fact, the straightest segment of the gesture trajectory does not correspond to any possible LOP.

Some parts of the gesture are performed many centimeters above the surface of the chart, while other are performed with the fingertip in contact with the chart. Real LOPs are drawn by putting a pencil in contact with the surface of the chart. Putting the fingertip in contact with the chart may add perceptual salience to these segments of the

gesture. The two strokes that correspond to the proposed LOPs are made with the fingertip in contact with the chart.

Finally, the two occurrences of the indexical "that" in BR's utterance are produced in synchrony with the two meaningful strokes and add to the perceptual salience of those strokes. These words mediate the allocation of attention, of speaker and listener, to the gestural performance. They highlight two moments of the performance, but the conceptual content of those moments and the relationship between them is not in the words alone; it is in the interpretation of the environmentally coupled gesture. Streeck (1993) says that gestures are "exposed" through coordination with indexicals in speech.

In summary, the imagination of proposed LOPs includes gestural strokes that pass near the depictions of landmarks and the estimated position. These are straight quick strokes, bookended by gestural holds, performed in contact with the surface of the chart and temporally simultaneous with highlighting utterance elements. These actions may be coordinated with eye movements tracing the LOP or saccading between the depictions of the landmarks and the EP.

The visual and somatosensory systems generate many representations of the location of the points of interest and the spatial relations among them. There are retina-centered representations as well as head-centered and body-centered representations. Each representational systems may encode multiple features such as location, direction of motion and velocity. Conceptual and visual constraints - represented deep in the MGC system concerning what a LOP can look like and where it can occur on a chart - support the imagination of possible lines of position.

Observational data cannot tell us in what order or precisely how the cues that distinguish virtual LOPs from other segments of the gestural stream are perceived, processed, or combined. The segmentation of the gesture stream into meaningful and non-meaningful passages depends on the predictions of the participants. Because they understand the activity, they are anticipating meaningful LOPs. When deep anticipation meets sensation, it generates perception of the relevant cues. Without committing to specific predicted cues, the relevance of the physical properties of the gestures can have a better or worse fit with the prediction of a possible LOP. Where cues fit those expectations, LOPs are perceived. Where cues do not fit those expectations, they may

instead fit the expectation of repositioning the hand, or of reflecting on where to go next.

It is surely a mistake to focus only on the parts of the gesture that represent LOPs. All of the gestural performance is interpretable by and meaningful to the gesturer as well as to the audience. As a viewer, I see provisional LOPs, but I also see repositioning of the hand and even indecision or reflection in pauses and circling the finger above the surface of the chart. Perhaps I even see stylistic flourishes in the velocity profile of the gestures. The predictions of the MGC system continuously respond to the sensory evidence to create meaningful perceptions. As is typical in human interaction, several projects are engaged in parallel, and the gestures can have different meanings in each of the projects. In the nominal task of choosing landmarks, a gesture may be interpreted as enacting the phenomenal object of interest, a LOP. In P's incidental task of monitoring BR's progress as a navigator, BR's momentary wandering gesture might be taken to mean that BR is unsure about the next LOP. In the supplementary task of constructing and maintaining social relationships, P's hand blocking BR's access to the plotting area could be seen as an assertion of P's authority over the process.

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

The simple acts of seeing the landmarks and the ship's estimated position on the chart bring visual process into coordination with structure in the chart and with memories for the depiction of the landmarks. This is already a complex process because the memory may be recall of specific depictions of known landmarks and/or recognition of landmarks through the interpretation of the graphical conventions used in cartography. In either case, these marks on the chart are recognized as depictions of landmarks and the previously plotted estimated position.

So why gesture? Superimposing gesture on the meaningfully interpreted chart surface a navigator activates predictions of motion in the visual system, and predictions of the trajectories of motion of the hand and fingers in the somatosensory system. The hands, guided by conceptually meaningful visual and motor predictions, act in the world thereby producing new richer, more complex, and more integrated constellations of

predictions. By acting and monitoring one's own action at the same time, one uses brain processes to guide activities that entrain even more brain processes. This is a self-organizing process that is located in the brain-body-world system.

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

The practice the navigators engage in is located in a complex cognitive ecosystem. The practice of gesturing to imagine lines of position brings into coordination many elements in a rich web of constraints that includes the tools of the job, the social relationships and division of labor among the people, the functional organization of the brain, and the culturally shaped ways of using the body. The high-level cognitive accomplishment, choosing appropriate landmarks, depends on all of these things. Each element of the system makes sense in the context of its relations to the other elements. This tight web of interrelationships is typical of real-world cognitive ecosystems.

The meaning of a complex emerges from the interactions among the modalities that include the body and material objects present in the setting. The effects of these interactions are not simply additive. A meaning complex may be built up incrementally or produced more or less whole, depending on the nature of the components and the relations among them (See Alač & Hutchins, 2004; Goodwin, 1994; Hutchins & Palen, 1997).

Typical interactions among humans are composed of many elements, the meanings of which emerge from the network of relations among the elements. For example, the representations of the provisional imagined LOPs are emergent properties of the complex activity system. Like components of a position fix, the parts of a meaningful human interaction only mean what they mean by virtue of the roles they play in the entire culturally constituted activity.

# 5.5. Navigation Discussion

In the few seconds of activity spanned by the navigation vignettes, the following phenomena appear:

**Locating bearing on hoey scale:** Multimodal engagement and sensorimotor contingencies; Continuous presence of and coupling to the world; Cultural conventions including number lines, numeric scales, and the 360-degree circle.

**Setting the location of the hoey arm**: Interpolation in action; Conceptual work in perceptual and motor processes deployed in the culturally constructed world; Alignment as a building block of many practices.

**Locking the hoey arm**: Thinking through tactile and motor modalities; Look ahead planning evident in the details of the grip on the hoey locking knob which affords rotation.

**Aligning hoey base with chart direction frame:** Combined chart space and body space (visual and proprioceptive modalities); Structure on the hoey base and on chart that afford alignment of the two direction frames using simple perceptual and motor processes; Another member of the family of practices based on alignment.

**Bringing the hoey arm to the LM depiction:** The articulation of chart space and body space gives meaning to body-based translational movement; The experience of meaning in perceptual and motor processes highlights the generative contribution of the MGC system; Meaning is made in the activation of the non-representational models of the hidden causes of sensation; Action is driven by the hidden causes of imagined sensation; Constraint satisfaction in action, constraints are supplied by depictions on chart (LM, direction frame) and the hoey (direction frame, angle of arm on the scale).

**Evaluating the relation of LM to EP:** Courses of action are trains of thought; Pausing motion provides a stable frame of reference for mutual adaptation or settling of constraining processes, prediction, sensation; The freezing of some modalities allows focus on those that continue to change. This is concentration; The dynamics of action, both real action and imagined action, are the dynamics of thought. The motor system has its own dynamic properties which may become constraints on thought in action.

**Detecting an error:** This is a spatial inference on a composite image that includes sensed and imagined structure; The anticipated track is "seen as" a sequence of positions; Translation without rotation has a clear meaning and a simple proprioceptive

signature; Embodied concepts; The thought IS the continuous dynamic experience of the body and the world.

**Correcting the error:** The proprioceptive signal of translation without rotation appears again; The dynamics of a movement stroke from beginning to end is characterized by smooth acceleration and deceleration indicating that the approximate target was already located; Thinking in action in meaningful space; Holding the pencil upright during translation shows that body state anticipates the next action which is also the next concept; Simultaneous control of multiple variables through the intrapersonal coordination of modalities.

**Checking the landmark identity:** The row and column table is an ancient cultural device for controlling the relationship between two variables; The bearing log is a long-term memory at the system level.

**Drawing the LOP segment across the projected track:** The LOP is drawn with a flourish as a sharp stroke (indicating finality of the process?); Are the dynamics of motion the dynamics of thought?

**The Three-Minute Rule:** Cultural processes collect partial solutions to frequently encountered problems. A serendipitous interaction between structures that developed for unrelated reasons produced a statistical regularity in the cognitive ecosystem that has been opportunistically captured by a cultural practice.

**Aha! Insight:** Complementary contents in intrapersonal modalities combine fortuitously to produce a novel insight. This is an illustration of a creative process that is underexplored.

**Choosing Landmarks:** Congruent contents in intrapersonal and interpersonal modalities produces stability that facilitates comparison of and reasoning over ephemeral representations.

The chart is a great example of dimensionality reduction. 3-D world reduced to 2-D. At large scale the spherical earth is rendered as a planar surface. This dimensionality reduction requires map projections. Every projection simplifies some computations and makes others impossible. On the Mercator projection in use on the bridge of the ship straight line is a line of constant bearing. This property of the chart fits with the practice of taking visual bearings and plotting them with a straight edge. The historical depth of the cultural legacy is apparent here. At small scale, the dimensionality

reduction appears as the 2-D rendering of the 3-D world of water surface with depth below and objects rising above sea level.

## 5.5.1. Imagining sequences: queuing and scales

The navigation examples made extensive use of numeric scales. As I pointed out in the analyses, the scale is a product of the projection of an imagined trajectory onto a real or imagined linear (or circular) array of spatial landmarks. It turns out there are many conceptual devices that rely on this same cognitive trick. One of the most familiar is the practice of queuing for service. In the following section, I begin with that practice and then try to radiate out to explore other members of the family of practices that are produced by projecting a trajector onto a spatial array.

In some cultural contexts, people seeking service arrange themselves in a queue as a way to control the sequence of access to services. People in a queue arrange their bodies in a linear array. The practice of queuing for service consists of three interlocking component practices. First, there is a cooperative social practice of forming linear arrangements of bodies. Second, there are spatial material (and perhaps architectural) practices that designate some location as the source of service. Third, there is a socially shared individual mental practice of seeing the linear arrangement of bodies with respect to the service location as a queue. These practices are mutually supportive and depend on one another for their meaning and their very raison d'être.

Let us focus for a moment on just the most obviously cognitive element of the queuing practice. Seeing a line as a queue is an example of the mapping of an imagined conceptual structure, what in cognitive grammar is called a trajector (Langacker, 1987, p. 231), onto a physical array. At the time, Langacker was working; there was no computational architecture that was well-suited to perform such a mapping. However, this is just the kind of thing that can be accomplished by a predictive processing computational framework.

The mapping of imagined structures onto one another produces what Gilles Fauconnier called a "conceptual blend" (Fauconnier & Turner, How We Think, 2002). In the case of the queue, an imagined structure is mapped onto a perceived structure. The perceived structure provides what I call a "material anchor" for the conceptual blend (Hutchins, 2005). This conceptual blend of a trajector with a spatial array gives rise to a particular

emergent property: a sequential ordering of the bodies of the individuals in the queue. The sequence of access to service is not present in either the physical line of people or in the trajector. It emerges only when some particular viewer blends the conceptual trajector with the perception of an appropriately ordered and situated physical array. Seeing the line as a queue is a cognitive practice because it makes possible a set of inferences. Who is next in line? Who arrived before whom? How far am I in space from (and how long must I wait before) getting service?

The social queuing practice is also cognitive, but for different reasons than the individual practice, and it implements different cognitive functions than the individual practice. The practice of queuing for service is, above all else, a cooperative public means to record and remember the order of arrival of clients. The queue also manages a forgetting function when people leave the line either before or after receiving service.

This everyday practice often takes place in complex social and institutional settings. Examining such settings can reveal the extent of the network of elements that are related in the cognitive ecosystem. In airports, for example, elaborate material arrangements serve to induce the formation of queues and shape the queues as they form. There may be patterns of lines painted on the floor, guide ropes or tapes, and signs such as "enter here" or "wait here for first available agent." The practice of forming a queue for service exists in a cultural ecosystem that includes services to be rendered (a set of facts about economic systems); the roles of service provider, who renders services, and client, who accesses services (facts about social organization); and locations in space at which service is rendered (facts about architecture).

Seeing an array of people as a queue integrates these elements in a particular way. It is an example of enacting a meaning by seeing the world in a particular way. A physical pattern that is open to many interpretations is "seen as" a particular, culturally meaningful, phenomenal object. The phenomenon of enacting meanings by "seeing" the world in particular ways (Stewart et al., 2010) is absolutely ubiquitous in human experience and is accomplished via cultural practices. When a line is being seen as a queue, other elements of the setting will be seen as instances of other roles in the queuing for service practice. This sort of fit suggests that cultural practices are composed of coherent constellations of mutually supportive component practices. In such a system, increasing the likelihood of any component increases the likelihood of the other components.

Forming a line from a group is a physical form of dimensionality reduction (Hutchins, 2012). A group of people occupying two dimensions of a surface approximates a one-dimensional array when they are seen as a queue. This dimensionality reduction does not take place in any person's mind. It takes place in the interaction of the MGC system with the space shared by the participants to the practice. The space typically includes features that bias the perceiving person in favor of seeing the group as a queue. Once the dimensionality reduction in the interaction with physical space has emerged, however, it supports or affords the cognitive practices of making inferences on the line "seen as" a queue. The experience of a one-dimensional line is more predictable than the experience of a two-dimensional crowd. The experience of a queue has lower entropy than the experience of a crowd. This increase in predictability is a feature of the relationship between an experienced MGC system and material structure created by a distributed system of cultural practices.

The superposition of a trajector on a spatial array of physical objects is the common constituent in a large family of cultural practices. In each instance, the projection of a trajector combines with different material and/or social practices to create composite practices. Each of these composite practices consists of a system of interlocking component practices. Each practice increases the predictability of experience by reducing the dimensionality of experience. Each one is located in a network of other practices and has mutually supportive relations to the other practices in the network. In the paragraphs below, I will mention practices and explore the local network of supporting elements for each.

Superposition of a trajectory onto visible or imagined objects is often applied to the natural world. Take, for example, the practice of seeing an array of stars as a constellation. One can see points of light with one's visual system. It takes a cultural practice to see a constellation. Decades ago, when I was trained in celestial navigation, I learned a number of strategies for attending to patterns in the night sky in order to identify specific useful stars. Many people know how to follow the so-called "pointer" stars in Ursa Major (also known as the Big Dipper) to find a useful star called Polaris. At the other end of the Dipper, the stars in the handle suggest an arc. One can "follow the arc to Arcturus and drive a spike into Spica." Calling the two stars on the lip of the Big Dipper the "pointer" stars and the mnemonic phrase oriented to the arc of stars in the handle of the Dipper are ways of speaking that help to organize the cultural practices of seeing. These discursive practices (Goodwin, 1994) exploit and activate the practice of imagining particular trajectors on particular visible arrays of points of light.

They are part of the local network in the cognitive ecosystem surrounding the practice of seeing constellations and using constellations to find particular stars.

In a similar way, the lexicon of sequence relations in general fits the practice of imagining a trajector on a spatial array. In the case of the queue, the ways of speaking about "first in line," "next," "back of the line," and so on are discursive practices that enter into relations of mutual reinforcement with the conception of the linear spatial array as a queue. Of course, these words and phrases are constrained by their use in other contexts, and these relations increase the tightness of the weave of the fabric of the cognitive ecosystem. The "tightness of the weave of the fabric of the cognitive ecosystem" depends on the pieces being learned. It's a tight weave when the component parts all mutually reinforce one another.

The memory practice known as the method of loci proceeds by associating an idea with each of an array of spatial landmarks and then imposing a trajector on the array of landmarks. This produces a sequential ordering on the set of ideas. It is also possible to imagine an array of objects and then project a trajector onto the imagined array. The phenomenon, known as fictive motion in language (Fauconnier, 1997; Langacker, 1987; Talmy, 1995), is produced by the projection of a trajector onto a real or imagined scene, as when one says, "the road runs down to the beach." The road is static, but this way of speaking adds a dynamic component to the experience of the static object. Fictive motion is a linguistic phenomenon, but employing it is not strictly a matter of knowledge of language. It is also a matter of knowing how to project a trajector.

The practices of numeracy and literacy also exploit the superposition of a trajector on real or imagined objects. Consider writing and reading. In the cognitive ecosystem, each of these practices constitutes the other's reason for being. How many such pairs are there in our modern cognitive ecosystem? The cultural practices of writing and reading assemble a complex configuration of resources, including writing implements, a physical text, physically inscribed in some medium, located in space, ways of moving the body (hands and eyes) with respect to the text, visual perception of words, extensive knowledge of language, etc. Writing and reading follow a conventional imposed trajectory (left to right and top to bottom for English).

Learning to read requires the domestication of visual attention (Goody, 1977). Similar processes of domestication of visual attention are at work in other sorts of reading, including reading natural phenomena such as the night sky, reading static

cultural notations such as those found in mathematics or music, and reading dynamic cultural representations, as in flying an airplane on instruments.Imposing a trajector on an array of written items creates a sequential list. In all of these activities, the domestication of visual attention produces culturally conventional trajectories of attention across spatial arrays of objects.

Dimensionality reduction is a key component of systems that must coordinate with visual attention. While an infinite number of scan patterns are possible, linear patterns are predictable and computationally inexpensive to describe. This is probably why linear trajectors are so common in the ecosystem.

The practices of reading and writing in coordination with the practice of printing lined pages produce yet another form of dimensionality reduction. Most writing systems include conventions for spatial layout under which the total surface area of a page can be seen as a single sequence of locations. The location of entries in a bound book may become part of a notation system. The apprehension of enduring relations among locations on different pages is possible only because the pages of the book have been bound together. Combining a numbered page sequence with results of the practices of reading and writing configures the total two-dimensional surface area of all the pages in a bound book into a single line of locations. If written entries are made in the book following the conventions of the writing system, then this single thread of locations throughout the book can be interpreted as a temporal sequence, which supports inferences about when the annotations were entered into the book. Because both books and cityscapes inherit numbering schemes from the same source, a location in a book with numbered pages is very much like an address of a building along a single street. Some of the practices for navigating a city are shared with the practices for navigating a book. For example, I'm looking at page 48 (or I'm on 48th Street) and I want to see something that I know is on page 53 (or 53rd Street). Which way do I turn the pages (or walk)? Relationships among practices in the cognitive ecosystem can create possibilities for the generalization of skill across activity systems.

The number line is a key component of many systems of mathematics. Projecting a trajector onto a linear array of number tokens creates a number sequence. Projecting a trajector on a linear array of numbers in order of magnitude and arranged with a constant interval produces a number line. This practice enables a variety of emergent inferences about relations among numbers. All sorts of scales for the measurement and expression of quantities are members of this family of cultural practices. For example,

when the numbers are read as times, a number line becomes a timeline. Reading an analog clock involves seeing a circle as a timeline. In this case, the trajectory is a curve rather than a straight line.

The practice of superimposing a trajector on a real or imagined array of objects is very productive, and there are many other practices in this family. A wide variety of other kinds of practices enter into symbiotic relations with members of the family: arrangements of bodies; architectural features; ways of speaking and verbal mnemonics; moral principles; arrangements of marks on surfaces; domesticated patterns of visual attention; the physical form of commercially produced writing surfaces; and more. Similarly, the members of the family are embedded in activity systems and participate in the accomplishment of a range of emergent distributed cognitive processes, including memory, representation of sequence, sequential computation, moral judgment, planning, time reckoning, arithmetic reasoning, navigation, search, and so on.

The variety of practices that are members of this family reminds us that cultural-cognitive ecosystems are heterogeneous and complex. As an object of study, these cognitive ecosystems fall into the cracks among the academic disciplines as they are currently organized. How many such families are there? There are certainly others based on conceptual primitives such as alignment and containment. Do such families have a place in the metaphysics of cognitive science?

## 5.5.2. Domesticated spaces and representation

At different moments in their work, navigators take two different interpretative stances with respect to the chart. Consider the navigator moving the hoey on the chart surface while seeking a coherent relation between a landmark depiction and the estimated position of the next fix. To what extent is the location of the hoey on the chart experienced as a location in the represented space (having meaning about locations in the world around the ship) and to what extent is it simply experienced as a location on the surface of the chart (in chart-table space) near where the plotting action will happen? The chart is sometimes a representation of something, it is not - the space around the ship. At other times, the chart is engaged as a thing-in-itself.

The chart is always meaningful, of course. It is always predicted to mean something. What people usually mean when they say that a chart is a representation is

that it depicts the space around the ship. In MGC terms, when this happens, the actual or imagined sensation of the chart is predicted in a network that sees the chart as the space around the ship. However, there are features on the chart that can be seen as representations of things other than the space around the ship. These elements depict features that do not exist in the natural world but that domesticate the representation of the wild world by combining it with cultural elements. The compass rose, and the lines of latitude and longitude are examples.

In the navigator's activity, the chart is most clearly a thing-in-itself when the hoey base is rotated to align the grid on the hoey with the grid (latitude or longitude) on the chart. This alignment is the satisfaction of a mechanical sub-goal that has no meaningful interpretation as an event in ship space. The chart is most clearly a representation of ship space when the location of a fix is compared to the location of the anticipated track or when the location of a fix is assessed with respect to the distance to an upcoming turn.

The balance between the salience of the two frames of reference shifts as task characteristics change. The chart will be a representation in some moments and a thing-in-itself in other moments. In some moments, perhaps it is both. There is no need for the interpretation of the chart as a representation of ship space and the interpretation of the chart as a thing-in-itself to be mutually exclusive. I imagine a probability distribution over the two conditions in which the chart could be interpreted either way to differing degrees.

To apprehend any material pattern as a representation of something other than itself is to engage in specific culturally shaped perceptual processes. When the chart is being perceived as a representation of ship space the activated parts of the MGC system include the features of the chart in the world, the flow of sensation produced by those features at the sensory surfaces, and the imagination or generation of the world in which the ship is located and through which it moves.

For example, in the moment when the plotter "sees" the displacement of the plotted position (taken as a location of the ship in the space around the ship) from the ink line of the anticipated track (seen as a potential path of the ship into the harbor outside with window) one could say that the anticipated track is being remembered. This memory takes part in spatial reasoning about how far the ship is away from where it should be at this point in time.

In the moment when the plotter notices that the position implied by plotting the Dive Tower landmark at a 057-degree bearing lies south of ("behind", given the northward motion of the ship) the last plotted position of the ship, one could say that the last plotted position of the ship is being remembered. And this memory enters a spatial reasoning process that concludes that the proposed LOP cannot possibly be correct.

External features, such as the ink line denoting the anticipated track and the little triangle denoting the last fix position, participate in remembering when their sensory traces constrain, bias, and activate those internal patterns that establish their roles in the current activity. This shaping of internal processes by external patterns/features is very important. The fact that humans put structure in their environments that can later do this work of shaping internal processes is one of the hallmarks of culture and distinguishes us from other species.

These examples illustrate the value of external representations as memories. It is not possible using internal resources alone to remember the chart and anticipated entry track accurately enough and, in enough detail, to judge this displacement.

In CitW, I noted that for the navigation system, a chart is a long-term memory for depicted features that are printed on the chart. It is an intermediate-term memory for the planned or anticipated entry track that is drawn with a pen in ink, and it is a short-term memory for the positions passed on this particular entry that are drawn in pencil. Note that the durability of the recording media matches the time scale of the use of the information encoded.

Features on the chart serve as each of these kinds of memory only in those moments in which they are incorporated in very particular ways into the activities of the team and, simultaneously, their sensory traces are shaping/shaped by internal MGC processes. The term of these memories (long, intermediate, short) is a property of the medium in the world in which the structure is expressed, print, ink, pencil. It is not a property of an MGC system, nor is it a property of any individual person.

While performing actions over and in contact with the meaningfully interpreted chart surface, a navigator activates predictions of motion to the visual system, and predictions of the trajectories of motion of the hand and fingers to the somatosensory system.

The notion of a "meaningfully interpreted chart surface ..." needs to be unpacked. This noun phrase implies a continuous dynamic coupling of the visual modality to the chart

surface (outward orientation) and (inward orientation) to a complex generative predictive processing (MGC) network. This network includes non-representational knowledge and a spotty memory of the chart. Memory can be spotty because the chart is present. The MGC network produces predictions of tuned sensorimotor contingencies. These networks change subtly as the chart is viewed, eventually predicting how the appearance of this particular chart will change with the movement of the eyes. This implies continuous dynamic coupling of proprioceptive modalities to the chart surface [outward orientation] and to additional MGC network structure. The developing representation of the specific contents of any particular chart is constructed on a foundation of knowledge of charting conventions, for example, the fact that north is away from the body (knowledge of the relationship of body-space to chart space to ship space). The spotty representation of the chart's contents may fill-in gradually as a consequence of implicit learning through interaction with the chart.

## 5.5.3. The Role of Friction in Action

There are three degrees of freedom to control in the plotting of an LOP: 1) The rotational location of the hoey arm with respect to the scale on the hoey base. Conceptually, this is the angular expression of the bearing of the landmark to be plotted with respect to true north. 2) The location of the chart depiction of the landmark. Conceptually, this is the location of the landmark in ship space. 3) The alignment of the hoey base with the directional frame of the chart. Conceptually, this links the bearing of the landmark as an angle with respect to true north to the cardinal directional frame of the chart.

These three degrees of freedom are all controlled via friction:

1. The hoey locking ring controls the amount of friction between the hoey arm and the hoey base. This friction constrains the rotation of the hoey arm with respect to the protractor scale. The hoey lock, when loose, permits free rotation of the arm around the protractor. When tightened, the lock produces friction between the hoey base and the hoey arm. This friction holds the hoey arm at a particular angle on the protractor. Of course, if sufficient force is applied, the arm can be made to rotate around the base even when the locking screw is tight. So, there is a balance here. Apply just the right amount of torque on the locking ring to produce friction that will hold the arm in place while the hoey is gently manipulated. Rough handling of the hoey will move the arm off the desired angle. At the level of the navigation team, the angle of the hoey arm is a very short-term memory for the bearing to the landmark. Friction determines the robustness of this memory. Friction also plays a role in the relationship of the fingers to the locking ring. The knurled surface of the lock ring provides friction for a good grip with fingertips.

2. The constraint of the location of the landmark is made available to the plotting computation by the pencil. The pencil is held in position, via friction of the pencil point on the chart paper, upon the depiction of the landmark. The friction of the pencil point on the chart surface holds the pencil in place as gentle lateral pressure is applied by the hoey arm to keep it in contact with the pencil. Too much downward force will crush the pencil tip, or, if the pencil is not aimed straight down into the paper, cause it to skid away from the target. Too little downward force on the pencil will allow the lateral push of the hoey arm to move the pencil point off the depiction of the landmark.

3. The grid on the hoey base is aligned with either the latitude or the longitude lines on the chart while gentle rotational (and or lateral) pressure keeps the hoey arm in contact with the pencil. This maintains the constraint of the landmark location while the hoey base is brought into alignment with the chart. Once the base is aligned, the hoey arm will lie along the bearing from the ship to the landmark. Pressure downward on the hoey base, pushing it into the surface of the chart, holds it in place. It is also held by the hand that is holding it, but it is more resistant to translational motion if friction with the chart is applied. Maintaining the lateral position of the hoey by controlling the lateral position of the hand is difficult and would require attention. Getting the hoey into position and then pressing down achieves the inhibition of lateral movement with a much coarser motor command. This amounts to exploiting friction to reduce the accuracy demands on the motor system. Since the LOP is now determined by the position of the hoey arm, the pencil can be removed from the depiction of the landmark and used to draw a short segment of the LOP in the vicinity of the expected position.

The control of these subtle frictional dynamics requires skilled, delicate, balanced, carefully sequenced application of force and torque. A little pressure now on the pencil and a slight torque on the entire hoey to keep the arm in contact with the pencil without either changing the angle on the hoey scale or skidding the pencil off the landmark are required. Light lateral pressure is applied to slide the hoey base away from and toward the body and left to right (not north, south, east, west) to align the grid on the hoey base with the direction frame of the chart (taken as a thing-in-itself). Then pressing down on the hoey base while simultaneously lifting the pencil tip from the chart. To bring the pencil to draw the segment of the LOP while maintaining downward pressure on the hoey base.

Clever use of friction simultaneously constrains these three variables. When thinking happens in the physical world, the properties of the world, including friction, are relevant resources for thought.

## 5.5.4. The Dynamics of Motion and Thought

In some activities, navigators allow their train of thought to be constrained by physical characteristics of the body and the motor system in particular. Movement is halted when

in reflection, slow when searching, and fast when moving to an identified new search location.

Thought can also be shaped by the constraints of gestural production. Pre- and post-stroke holds highlight or frame meaningful segments of a continuous gestural stream. It is not that the thought causes the holds, or that the holds (driven by independent constraints of gesture production) cause the organization of the thought. The thought is the highlighted stroke (together with the predictive processing network activation) bookended by holds.

The activity of plotting position is thinking in embodied action. The activity has properties that emerge in the continuous**,** embodied, multimodal engagement of the plotter with the chart via the plotting tools. The plotter, plus hoey, plus chart is a system. Continuously enacting (generating/predicting/anticipating) meaningful experience in multiple sensory modalities.

## 5.5.5. Can an Obsolete Form of Navigation be Representative of Other Activity?

Navigation is a special domain of activity, and this sometimes gives rise to concerns about the generalization of the findings made here. Navigation is unusual, and the kind of navigation described here is obsolete. Fortunately, navigation does not involve cognitive processes that are alien to everyday life. Rather, what is special about this setting is how well it supports enacted reasoning. The generalization of results must be tied to the distributions of the modes of thinking in the cognitive ecosystem, not to the specific characteristics of the setting.

# 6. Aviation vignettes

While the ship navigation vignettes focus on relationships among modalities in a single person and illustrate principles of embodied cognition, the aviation vignettes move the focus to relationships among modalities that span two or more actors. These vignettes illustrate principles of embodied interaction.

# 6.1. Fuel Leak

[10] A leak in the fuel system of an airplane is a very serious problem. In the best case, the amount of fuel available to run the engines could be reduced, calling for a diversion to an alternate airport short of the planned destination. In the worst case, a leak could lead to the destruction of the airplane in flight.

In the late 1980s, my team was given access to video recordings made in high-fidelity flight simulators at NASA Ames Research Center. In one of the simulated flights, a three-pilot crew confronted a fuel leak in a Boeing 727. The crew members were recruited from a major airline and were, in those days, flying the 727 in revenue service.

In the 727, the Captain (C) and First Officer (FO) sit facing forward at the flight controls and take turns with the duties of flying the airplane, navigating, and communicating with air traffic control. The Second Officer (SO, also known as the flight engineer) sits behind the First Officer. His seat can swivel to face forward, but most of the time, he faces the flight engineer's panel on the right side of the cockpit, where he monitors and controls the mechanical systems: engines, fuel, hydraulic, electrical, pressurization, etc. The fuel control sub-panel dominates the most accessible area of the engineer's panel.

The event analyzed below highlights the fluid transition between seeing the fuel control panel as a structure in the cockpit - a thing-in-itself - to seeing it as a representation of the fuel system, and back to seeing it as a panel. As with the navigation chart, the question is not 'What is a representation?' or 'What features do external representations have?', rather the correct question is 'WHEN is a structure a representation?'. Under what conditions should we say that an external structure, pattern, or process functions as a representation? In this vignette we see that the expectation of meaning, grounded in expert knowledge, applied to gesture, talk, and particular objects in the cockpit, places the experience of the fuel panel in a complex network of conceptual relations sometimes in the role of the fuel system, such that gestures and talk refer to events in the fuel system (panel qua representation), and sometimes in the role of a display in the cockpit that can be manipulated without reference to events occurring in the fuel system.

The physical layout of the fuel panel and its relations to previously encountered representations of the fuel system permit the crew to see the panel as an object in itself

and as the fuel system it represents. This allows the gestures performed over the panel to be interpreted as actions taken on the panel, or as events in the fuel system, or both. The speech is used in part to manage relations that are not easily expressed in gesture and also to move from one interpretive mode to another.

## 6.1.1. Detecting a Fuel Leak

The 727 has three engines and three fuel tanks. Normally, fuel is fed from each tank to its corresponding engine. About a minute before the segment analyzed below, SO noticed an unusually low reading on the fuel tank number three quantity gauge. Scanning the panel, he paused at the tank three gauge, momentarily continued the scan, and then snapped his head (and we presume eyes) back to the tank three gauge.

With any remote sensing system, there is always a question of whether what is observed is really the behavior of the system or if it is a malfunction of the sensor display. When a fuel leak is suspected, a flight engineer should press the fuel quantity test switch to confirm that the fuel quantity gauges themselves are operating properly. Pressing this button moves all the fuel tank gauge needles simultaneously to different positions to test for responsiveness. When the fuel quantity test button is released, the needles return to their previous positions, where they, hopefully, correctly indicate the levels of fuel in the three tanks. As expected, SO checked the operation of the fuel gauges by pressing the fuel quantity test switch. Figure 6.1.2 shows a diagram of the fuel panel. The fuel quantity test switch is circled in red at the upper left corner of the panel.

Once SO has confirmed that all the gauges are working properly, the next step is to locate the fuel leak. A leak could be in one of two places: in the tank itself or somewhere in the fuel line between the tank and the engine. To determine if the leak is in the tank, the fuel must be isolated in the tank by turning off the boost pumps in that tank. If the gauge still indicates a decline, then the leak is in that tank. If there is no decline in fuel quantity when the boost pumps are turned off, the tank is sound, and the leak must be somewhere in the fuel line. This is an even more dangerous situation than a leak in the tank, because the fuel may be escaping into the fuselage where it could ignite and destroy the airplane. Thus, the next diagnostic action is to turn off the boost pumps that feed fuel from tank three to engine three.

The goal of turning off the boost pumps for tank three cannot be accomplished directly, however, because doing so would cause engine three to flame-out (quit running). Before the boost pumps can be turned off, SO must first establish an alternate fuel supply for engine three. Once this has been accomplished, the pumps can safely be turned off and the gauge monitored for further fuel loss.

SO configured the cross-feed valves to establish fuel flow from tank one to both engines one and three. He then turned off the boost pumps that drive fuel from tank three to engine three. Even with the pumps turned off, the rate of decline of fuel quantity in tank three was noticeably greater than from the other tanks. This meant that the leak was in the tank, which was bad, rather than in the fuel lines inside the fuselage, which could have been catastrophic.

## 6.1.2. Reporting the Fuel Leak Diagnosis

After completing his diagnosis of the fuel leak situation, SO turned in his seat to face the front of the airplane while addressing the captain and FO. No gestures other than body orientation accompanied this announcement.

12.00.43 SO: *well it looks huh like a funny situation. we have a fuel leak or something \2\ in number three tank*

A salient part of a pilot's understanding of a fuel leak is that it is a situation that must be dealt with quickly. In response to the SO's announcement, the captain and FO turned in their seats to face the SO and the engineer's panel (Figure 6.1.1).

**Figure 6.1.1. Attending to the fuel panel**

*Frame from a video courtesy of NASA Ames Research Center*

After SO's announcement, the crew members collectively knew what SO suspected (a fuel leak) and where he thought the problem was located (in fuel tank three). With that information, the crew members prepared to attend to the problem.

Once the captain and FO were situated, SO began his explanation of the problem without further prompting. As SO spoke, he turned his seat to face the fuel control panel. Because the SO uses the panel to explain his actions, it is useful to get familiar with the relevant properties of the panel.

**Figure 6.1.2. Fuel Panel Diagram**
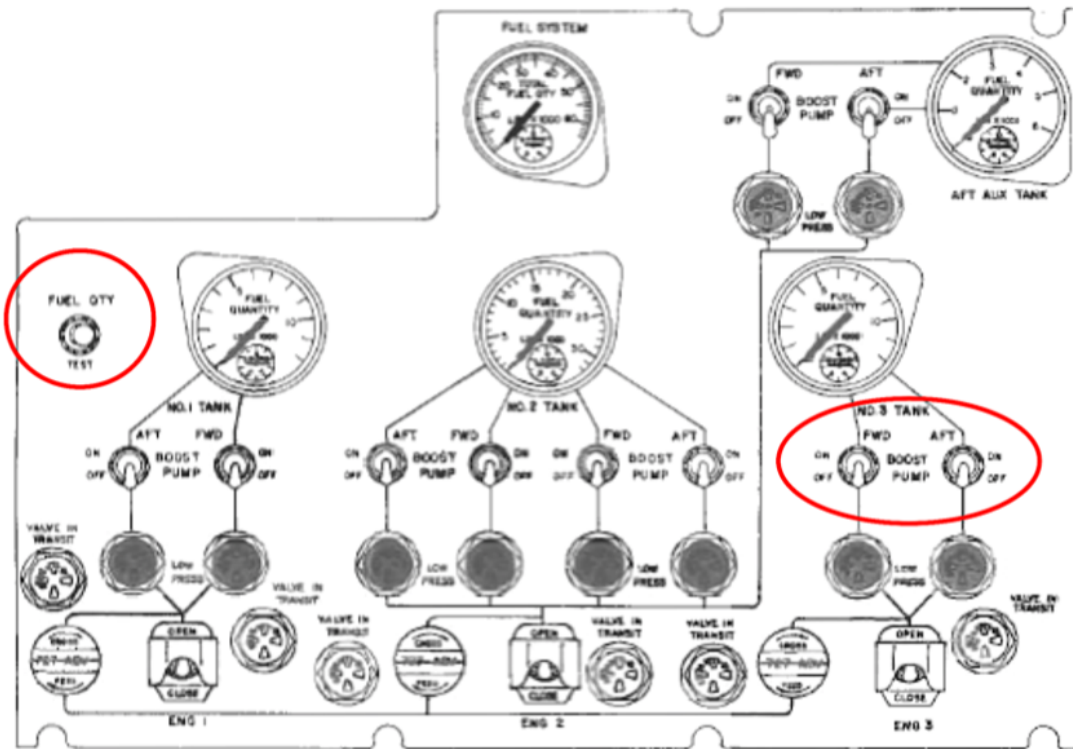
*Fuel Quantity Test Switch (upper left) and Tank Three Boost Pump Switches (lower right)*

The spatial layout of the panel (Figure 6.1.2. provides a diagram taken from a 727-training manual) is topologically, but not metrically, identical to the spatial layout of the fuel system that it depicts. The topological relations among panel components (e.g., the quantity gauges, painted lines, and pump control switches) are the same as the topological relations among the system components (e.g., fuel tanks, fuel lines, and pumps). Components that are higher on the panel generally correspond to fuel system components that are forward in the airplane. Components that are to the right on the panel generally correspond to fuel system components that are on the right side of the airplane. The panel is simplified by omitting depictions of check valves and other components that cannot be controlled from the panel.

The topology of the painted lines and switch positions enables the crew to make conceptual inferences with simple and robust perceptual skills. For example, figuring out where fuel will flow can be accomplished by moving visual attention along lines on the panel.

**Table 6.1. Correspondences between components on the fuel panel and components in the fuel system**

| In the system | On the panel |
|---|---|
| Fuel tank | Quantity gauge |
| Fuel line | Painted line |
| Boost Pump | Pump control switch |
| Pressure sensor | Pressure indicator |
| Valve | Valve control switch |
| Fuel flow established by valve position | Position of valve control with respect to painted lines |

The simplified topology of the panel permits the pilots to reason about the state and behavior of the fuel system by "seeing" the panel in a particular way. For example, the valve controller has a line painted on its top surface. When the controller is in the cross-feed position, this line appears to connect the painted lines that depict the fuel line arriving at and departing from the valve. The rotational action of the cross-feed valve controllers, combined with the shape of the controller knob, facilitates seeing the open and closed positions of the valve as flow through or blocked flow. These may seem to be trivial design features, but they have important cognitive consequences.

The fuel system itself is a collection of physical components distributed through the wings and fuselage. The system as a whole cannot be seen from any real vantage point, but the pilots can "see" the fuel system by seeing through the fuel panel. In fact, only through seeing fuel panels and diagrams such as Figure 6.1.2 do pilots have any experience of the topology of the fuel system. As with any material structure that can be seen as a representation, it is possible to see either the structure itself or to see the thing that is represented. Sometimes it is possible to see both at once. Understanding SO's performance requires several shifts in seeing. How do gesture and speech guide these shifts between the perceptual stance in which the panel is seen as a thing in itself and the perceptual stance in which the panel is seen as a representation of the fuel system?

With C and FO oriented to him and his panel, SO continued his account of the problem.

*SO: I don't know we must be losing it very quickly you see right now I-\2\I turned the pumps off ok I tried to feed from number one to both engine one and three but we're still losing in number three quite a bit*

| SPEECH | GESTURE IN SPACE |
|---|---|
| right now | Placed index finger on, but did not depress the fuel quantity test switch |

SO began his account by gesturing to (placing his finger on but not depressing) the fuel quantity test switch while saying "right now." There was nothing in SO's words about the fuel quantity test button. He had pressed it earlier, of course, as part of the diagnosis procedure. But during his explanation of the problem to the other crew members, SO neither mentioned nor pressed this button. He only touched it. We believe that the other crew members interpreted this as an indication that SO had already tested the gauges (in fact, he had).

In the MGC framework, C and FO would have experienced, via motor resonance, an activation of the motor imagination of pressing the button. The fuel quantity test button is spring loaded and light pressure is required to depress it. It's a momentary switch. While holding it down, one watches the gauges to verify that the needles on the gauges move freely. The button is smooth black plastic, about six millimeters in diameter, with a slightly concave head. Does all of this come to mind for C and FO when SO touches the button? Mirror neurons allow them to follow along with motor imagination as they watch the SO's actions. This gives them an additional resource to create a richer imagination of the concepts SO is presenting. This richer imagination can support their joint problem-solving.

The joint work of dealing with an in-flight emergency gives rise to a sort of contagion of expectationthat is grounded in a history of shared experiences. SO's actions give rise to motor activations in the other pilots, bringing them to imagine the SO's conceptual project of determining the location of the fuel leak. This contagion is carried by gesture and experience of the shared world as well as by speech. The other pilots can imagine what comes next. Their MGC systems are entrained by communication.

Notice that the fuel quantity test button does not appear in the table of correspondences between components on the fuel panel and components in the fuel system. The fuel quantity test switch differs from all other elements of the panel. All the other elements can, under some conditions, be interpreted as being "about" objects in the fuel system, but the fuel quantity test switch is strictly "about" a set of components, the quantity gauges, on the panel.

Thus, touching the fuel quantity test button must be experienced as being about the panel as a thing-in-itself rather than about the fuel system. In order for C and FO to interpret SO's gesture to the fuel quantity test switch, more than a shared understanding of its function was necessary. It was not enough that they all had a similar model of the switch's function. They needed to know that the others had a similar model of the function as well. This kind of intersubjectivity underlies all of the meaningful actions on the panel.

The words "right now" indicate that the actions being reported took place recently. SO did not delay in notifying C and FO about the problem. These words also gave a sense of immediacy to the situation. They place something in the present time, but what is not yet clear. The speech and the gesture are working independently of each other here, each conveying information about different aspects of the same conceptual project: speech about when the actions took place, gesture about the initial action taken.

| SPEECH | GESTURE IN SPACE |
|---|---|
| I -\2\ I turned the pumps off OK | Brought index and middle fingers down onto the tank three boost pump switches |

SO next made a motion over the number three tank boost pump switches that mimicked the motion used to turn the pumps off. The switches were already in the off position. The combination of the gesture and the state of the panel, and the knowledge that boost pumps are normally on in flight, made this action unambiguous. The past-tense word "turned" provides temporal information that cannot easily be conveyed with gesture. Gestures are always in the present tense. Simultaneously, the gesture specifies which pumps have been turned off, a specification that is lacking from the speech. Speech marked the gestures as a reenactment of what SO had already done. The verbal statement did not indicate which pumps had been turned off, but the fingers did. The

location of the gesture in the space of the fuel panel thus resolved an ambiguous reference in the verbal stream. The verbal component provided temporal markings that were lacking from the gesture, and the gesture provided aspects of indexical reference that were ambiguous in SO's words.

The boost pumps being off raises the question of where the fuel for engine three was coming from. The topology of the panel facilitates certain inferences about the functional behavior of the fuel system, and SO next moved to demonstrate these inferences to the other crew members.

| SPEECH | GESTURE IN SPACE |
|--------|------------------|
| I tried | Moved fingers from the tank three boost pump switches to the area of the tank one boost pump switches |

SO changed topics (introduced a new conceptual object) at this point, and his gesture directed attention to the other side of the fuel panel, where subsequent events would be described. He was now beginning to explain how he established an alternative fuel source for engine number three. The use of the past tense "tried" placed the action referred to in the past with respect to the present course of action.

| SPEECH | GESTURE IN SPACE |
|--------|------------------|
| to feed from number one | Moved index and middle fingers up and down between the tank one quantity gauge and the boost pump low pressure indicator lights |

Here, the gesture and the speech were almost completely congruent. The gestures draw visual and motor attention to the lines painted on the panel that depict the pipes in the system that move fuel from the number one tank, through the boost pumps, and to the engine's one fuel feed valve. The gestures highlight specific elements of the panel, but the speech "feed" refers to an event in the fuel system. In this moment, the panel is "seen as" the fuel system. The up and down motion of the fingers is seen as the movement of fuel through the pipes. The panel is, at this moment, a representation of the fuel system.

| SPEECH | GESTURE IN SPACE |
|---|---|
| to both | Hand raised away from the surface of the panel |
| engine one | Moved hand across panel to the area of the controls for tank and engine three, pointing to engine three cross-feed valve control |
| and three | Moved hand back across the panel to the area on the controls for tank and engine number one |

In the brief statement, "I tried to feed from number one to both engine one and three," SO explained that he had remembered to feed fuel to engine three before he turned the tank three boost pumps off. The gesture accompanying this section was complex and quickly executed. SO pointed to the tank one gauge, to the tank one pump switches (which were in the ON position), then to the engine three cross-feed valve controller, and to the engine one cross-feed controller. These gestures drew attention to the controllers that indicate that the valves were open and supplying fuel to engines one and three from tank one. Some of the motions of the hand also followed the flow of fuel through the system. The panel continues as a representation of the fuel system.

| SPEECH | GESTURE IN SPACE |
|---|---|
| but | Pointed with index finger to the engine three fuel gauge |

Having established the alternate source of fuel for engine three, SO pointed to the engine three fuel gauge. This was the locus of the problem. SO marked with a gesture a return to the topic of the fuel level in tank three. The logical disjunction, "but," marks a violation of a possible conceptual continuation of the situation described before. The elements that stand in disjunction are not yet clear but will be made clear by what follows.

| SPEECH | GESTURE IN SPACE |
|--------|------------------|
| we're | Flicked the face of the engine three fuel gauge with middle finger |

SO flicked the gauge with his finger. This is a common technique among pilots to free a gauge needle that is believed to be stuck. From a strictly functional point of view, this is a useless action. SO detected the fuel leak by observing the rapid movement of this fuel gauge needle. The fact that it was possible to detect the fuel leak this way is evidence that the needle is not stuck. Furthermore, he had tested the freedom of the needle movement with the fuel quantity test button. Thus, the flick cannot be an attempt to free a stuck needle.

This flick was not performed in SO's original diagnosis and was not a report of a previous action. Rather, it was a new action performed while the other crew members looked on. The use of the first-person plural "we" indicates that the problem is shared by the crew. The present tense "are" returns the narrative from an account of earlier actions to the current situation. Flicking the gauge breaks the experience of the panel as a representation of the fuel system. The flicking has no meaning in the fuel system, but it does have meaning on the panel experienced as a thing-in-itself. I have even seen pilots flick gauges on digital displays, which, of course, cannot affect the gauge reading.

Because this action cannot be functional, we might ask what other kind of role it might be playing here. For one thing, it returns the narrative to the present. It is a way of emphasizing that the fuel level shown by the number three tank gauge is the salient problem. At a more abstract level of description, flicking a gauge is a way to produce an expected reading when an unexpected reading has been encountered. In that sense, this action could also be read as an assertion by SO that he would have liked the behavior of the gauge to be other than it was.

| SPEECH | GESTURE IN SPACE |
|--------|------------------|
| still losing | Repeated jabbing motions at the face of the tank three gauge with the index finger |

SO then emphatically gestured to the tank number three gauge, while he said, "still losing." This last gesture drew attention away from the function of the needle (that which the prior gesture, the flick, demanded) to the actual fuel quantity level that the needle was indicating. In this moment, the gauge is again a representation of a feature of the fuel system.

| SPEECH | GESTURE IN SPACE |
|---|---|
| in number three quite a bit | Moved hand away from the panel and into lap |

Finally, SO returned his hands to his lap, indicating that his turn was completed.


## 6.1.3. Summary

The diagnosis of a fuel leak is a conceptual project with many parts. In some utterances, the talk and gesture modalities have congruent contents. In others, the contents are complementary, as when SO's talk refers to "the pumps" and the gesture indicates which pumps. Some utterances treat the fuel panel as a representation of the fuel system, while others treat the same panel as a thing-in-itself.

The shifts from the interpretation of the fuel panel as a representation of the fuel system to the interpretation of the fuel panel as a thing-in-itself seem to be constrained by a requirement for agreement between the interpretation of the fuel panel and the informed interpretation of the meanings of gestures and talk. The default expectation in the flight deck setting must be that the fuel panel is a representation of the fuel system. Where gestures and talk can be interpreted as being about objects and events in the fuel system, the panel is interpreted as a representation of the fuel system. The whole is formed through simultaneous expectations about three semiotic resources here. The network's generation of expectations about the representational status of the fuel panel, the network's generated expectations about the spoken words (their meaning), and the expectations about the gestures in visual and motor modalities (their meaning) all mutually elaborate one another - which means that the various parts of the network impose constraints on one another and respond to constraints imposed by one another. The network adapts to simultaneously satisfy these constraints. The network configuration in which these constraints are satisfied is the meaning of the event.

In the analysis of this vignette, I suggested that the members of the crew might share an embodied experience of the operation of the fuel system. In the next vignettes, we will see strong evidence about the richness of the shared imagined workplace.

# 6.2. Approach to Stall Recovery

[11] This section examines the interaction among a Boeing flight instructor and two Japanese pilots who are transitioning from other Boeing airplanes to a particular model of the B737. We are especially interested in cases where a gesture or other non-verbal element produced by one participant stands in a relation of mutual elaboration with a spoken element produced by another participant. We focus on this because such instances could not occur if the participants did not share a sense of the conceptual project being developed. These are moments in which parts of an otherwise invisible shared conceptual network become visible. The participants use a variety of semiotic resources to bring to the fore selected aspects (but rarely the whole) of a conceptual structure.

In earlier vignettes, I introduced the idea of the semantic relations among the contents of modalities in a multimodal system. Such contents may be congruent (gesture and talk refer to the same aspect of a complex concept) or complementary (gesture and talk refer to different aspects of a single complex concept). When the modalities of interest occur in different individuals in interaction with one another, the contents may be produced at different times. Potential temporal relations among the appearance of modality contents are concurrent (occurring at the same time) and offset (occurring at a remove in time).

This entire event is a powerful illustration of the richness of the pilots' shared models of the world. The elements of talk and gesture that are produced in interaction are not just evidence of the underlying models; they are parts of distributed thought processes. There exists a single shared complex model of the behavior of the airplane. The participants take turns (but not really turns, sometimes working in parallel rather than sequence), highlighting or activating elements of that model.

As part of a research agreement with Boeing's Flight Deck Concept Center under the direction of Barbara Holder, my team made video and audio recordings of training activities in which pilots from non-US airlines transitioned to the Boeing 737NG. We recorded instructor/pilot interactions both in the simulator and in the briefing room

before and after simulator sessions. Saeko Nomura-Baird recorded a total of thirty-seven hours of training for three Japanese pilots at Boeing's training center in Seattle, Washington. This training was conducted in English. In what follows, Saeko and I analyze two minutes and thirty-seven seconds of interaction between an American instructor and two senior Japanese pilots engaged in a pre-simulator session briefing. Both of the pilots were already qualified to fly as captains in other models of Boeing airplanes.

One of the maneuvers practiced by the pilots is called an Approach to Stall Recovery. An airplane stalls when the smooth flow of air over the wings separates from the upper surface of the wing. When this happens, the wing ceases to produce lift, and the airplane falls. This is an extremely dangerous situation. Pilots never practice taking an airliner into a full stall, even in the simulator. Instead, they practice responding to the first indications of an impending stall; thus, the approach to stall recovery.

There is a difference between the way Boeing teaches this maneuver and the way it is taught and practiced at the airline for which the pilots work. A pilot can approach a stall by holding back pressure on the yoke as the airplane decelerates. To recover from a stall approached this way, a pilot increases engine power and simply relaxes the back pressure on the yoke. This is how the maneuver is taught by the airline for which the pilots work. This is a way to produce an approach to a stall, and it allows pilots to experience and understand the cues that indicate that a stall is imminent. However, stalls in real-world operations are seldom approached by pilots holding back pressure on the yoke. A more realistic way to approach a stall is to use stabilizer trim to neutralize control yoke pressures while decelerating. To recover from a stall approached this way, the pilot adds power and must push the yoke forward to prevent the nose of the airplane from rising. If the nose is allowed to rise, it can cause the airplane to stall again. Quite a lot of force is required to push the yoke forward in this situation. While pushing, the pilot applies nose-down stabilizer trim so that a flying attitude can eventually be maintained without continuing to push the yoke forward.

**Figure 6.2.1 FCOM The "Approach to Stall Recovery" procedure as it appears in the Flight Crew Operations Manual**

| Maneuvers | Chapter MAN |
|---|---|
| Non-Normal Maneuvers | Section 1 |

## Approach to Stall Recovery

The following is immediately accomplished at the first indication of stall buffet or stick shaker.

| Pilot Flying | Pilot Monitoring |
|---|---|
| • Advance thrust levers to maximum thrust.<br>• Smoothly adjusting pitch attitude* to avoid ground contact or obstacles.<br>• Level the wings (do not change flaps or landing gear configuration).<br>• Retract the speedbrakes. | • Verify maximum thrust.<br>• Monitor altitude and airspeed.<br>• Call out any trend toward terrain contact. |
| When ground contact is no longer a factor:<br><br>• Adjust pitch attitude to accelerate while minimizing altitude loss.<br>• Return to speed appropriate for the configuration. | |

**Note:** *At high altitudes it may be necessary to decrease pitch attitude below the horizon to achieve acceleration.

Boeing teaches the maneuver using this second, more difficult, approach to stall recovery technique. The pilots in these vignettes refer to this as the "Boeing way." The procedure shown in the Flight Crew Operating Manual (FCOM) (Figure 6.2.1) is a generic procedure that can be used to recover from an approach to a stall in any configuration (setting of flaps and landing gear). The pilots are preparing to practice an approach to a particular kind of stall event known as a departure stall. This is flown with the landing gear retracted, the flaps extended at 5 degrees, and with a 20-degree bank attitude (Figure 6.2.2). The flap setting is a key element because it determines the speed at which the maneuver is begun as well as the target speed for its completion.

**Figure 6.2.2. The practice procedure for the recovery from an approach to the "departure stall" as it appears on the computer monitor in the pre-simulator briefing**

### #1: FLAPS 5

### 20° Bank

- Start: Flaps 5. FLAPS 5 speed

- N1 45% (approx 1 knot/sec decel)

- Establish 20° bank  (PF check VSI, ALT, PLI)

- Stick shaker

    o   Smoothly apply MAX Power (PM adjust to GA)

    o   Level wings, do not change config, retract SB

    o   Power comes up, apply nose down trim

    o   Airspeed increases, lower pitch to 5– 6°

    o   Approaching Flaps 5 speed– 65%

- Finish: FLAPS 5 speed

Because coherent meaning structures are created by multiple utterances, we organize the presentation by cases rather than by utterances. A single conceptual object is created in each case. Each case is given a brief descriptive title and begins with a concise description of the conceptual object that is constructed in the case.

Excerpts from the transcript are provided with each case. In the transcripts, the two pilots are referred to as PF (Pilot Flying) and PM (Pilot Monitoring) to reflect contemporary usage. The instructor is designated by the letter I. Punctuation is used to represent intonation: A period indicates falling pitch, a question mark rising pitch, and a comma falling contour, as would be found, for example, after a non-terminal item in a

list. A colon indicates lengthening of the current sound. Numbers within single parentheses mark silences in seconds and tenths of a second. Words within parentheses indicate uncertain transcription. Underlining denotes words that are spoken in synchrony with gestures.

## 6.2.1. Flaps 5 Speed

Conceptual object: Airplane dynamics; accelerate an airplane from stall speed to flaps 5 speed, constructed from the point of view of the crew.

**Figure 6.2.3. Flaps 5 speed**



The instructor read the procedure from the FCOM, tracing the text with his left index finger as he read, "Return to speed appropriate for the configuration." He then looked at the computer monitor and pointed to highlight the portion of the procedure described by the words "Finish: FLAPS 5 speed." The instructor elaborated this part of the maneuver, and as he withdrew his right hand from a full-hand point to the procedure shown on the computer monitor, he said, "flaps five speed." Simultaneously, PF positioned his hands as if holding the control yoke and pushed them forward (see Figure 6.2.3). This gesture enacted the control input needed to return to flaps 5 speed. Notice that the instructor's utterance does not specify the sort of control input that will be needed to return to flaps 5 speed. PF knows that in order to accelerate, he will have to push the yoke forward. Thus, the gesture provides the cause that is not present in the verbal description of the effect "go to flaps five, flaps five speed." The words provide the effect that is not present in the gesture. The two elements mutually elaborate each other as a metonymic cause-and-effect relationship. The idea of returning to flaps 5 speed gives meaning to the push gesture, and the push gesture specifies an element of the process of returning to flaps 5 speed. This is a collaboratively constructed multimodal utterance in which the

instructor's speech and the pilot's gesture are temporally concurrent. This means that the pilot's and the instructor's imaginations are running through time in parallel. The instructor's words and the pilot's gesture are semantically complementary; each highlights or brings into imagination a different but related aspect of a single conceptual object. Semantic complementarity across participant constructions is strong evidence of the existence of the shared conceptual object.

In the MGC framework, the fact that the pilots share a conceptual model means that they share a model of the hidden causes of the sensations of talking, reading, and gesturing about the concepts. It does not mean that they share the same networks or the same patterns of connections among elements. Rather, it means that however their networks may be arranged, they share the functional ability to make the same predictions in the same contexts.

The cognitive ecosystem of the pre-simulator briefing suggests another cognitive function for PF's gesture. Considering that PF was representing a component of the procedure that he expects to execute in the simulator, it might also be a sort of pre-enactment that could facilitate memory for the procedure later.

## 6.2.2. Back Pressure Only

Conceptual object: To decelerate an airplane in level flight, reduce power, and hold back-pressure on the yoke, constructed from the PF's character viewpoint.

PM begins this project by saying, "Yes, I know difference between Boeing and [company X]'s procedure. Our procedure just trim out at flap five speed." The instructor provides a verbal continuation, saying "and then." This, by itself, is a collaboratively constructed verbal utterance. But there's more. The instructor also simultaneously gestured to model pulling back on the yoke (Figure 6.2.4). The added gesture here fills in content for the projection into the future of his own words, "and then," and projects a syntactic structure for PM to complete verbally.

Every pilot knows that if you do not trim, in order to decelerate, you must pull back on the yoke. The instructor used that knowledge to anticipate the projection of PM's words with his gesture. This projection was especially well marked as PM had stated that his company's procedure is different from Boeing's procedure, which they had discussed, and which involves trimming to the stall speed. Thus, the instructor's gesture is

semantically congruent with and temporally precedes PM's spoken words, "back pressure." The gesture also has a relation of mutual elaboration to the concurrently produced words "and then." The semantic relation here is complementary (synecdoche), because the talk represents a sequence in which the back pressure enacted in gesture is a component action. The initiation of the instructor's gesture was anticipatory, but he held it while PM continued speaking, saying, "keep back pressure only." By the end of this statement, the instructor's gesture and PM's speech were semantically congruent and temporally concurrent. This feels a bit like a jazz improvisation where the musicians anticipate one another and explore complementary musical elements, finally returning together to the tonic.

**Figure 6.2.4. Back pressure only**



PM: Yes, I know difference between Boeing and (Company X)'s procedure. Our procedure just trim out at flap five speed (0.2)

I: and then
[makes two fists to represent holding yoke and pulls toward his chest]

PM: keep back pressure only
[I continues holding his two fists near his chest]
not applying any more trim.

The conceptual projects developed by the pilots and instructor are richer and more complex than any turn of talk or gesture can depict. Talk and gesture highlight components of the underlying project but never approach an exhaustive account. Sharing the underlying model, the pilots experience a flow of expectations shaped by their own trains of thought around the network of concepts in the model and the cues provided by the actions and words of the others.

## 6.2.3. It's Realistic the Boeing Way

Conceptual object: A comparison of techniques, from two implied character viewpoints. Both assume a pilot seated in the flight deck. PM takes a viewpoint of looking down on the stabilizer trim indicator, while the instructor assumes a pilot flying from the right seat holding the right horn of the control yoke.

**Figure 6.2.5. Realistic the Boeing way**

PM: But a, it's realistic the Boeing way. (0.5) Because, ah

[*I makes trim gesture*]
 PF: Hmmm.

PM: We always manage to keep our trim...
[*I nods continuously*]
I: Uh, huh.

PM: ( ) you know, forward out of habit.
[*I nods continuously*]



By using the words "the Boeing way," PM refers to the earlier discussion that established the contrast between his company's technique of entering a stall recovery maneuver using back pressure on the yoke only (no stabilizer trim) and the "Boeing way" that involves trimming as the airplane slows on the stall entry. Framing the topic as "It's realistic the Boeing way" constructs an implicit comparison between the techniques. The instructor knew this, and the movement of his right thumb models the action that PM would take as a pilot flying (from the right seat) when he trims the airplane. The conceptual schema was clear at the pause before the word "Because." Possible projections included elaborating on either the realistic or the not realistic method. The instructor's gesture is an iconic representation of an anticipated spoken description of trimming in accordance with the realistic method. But PM does not follow that projection of the schema. Instead, he elaborates on the not-realistic method that his company uses.

This case is interesting because the instructor's gesture is semantically congruent with an anticipated spoken representation that never occurred. We could even say that the instructor's gesture is positioned and formed to facilitate the production of a verbal element with which it could be both temporally concurrent and semantically congruent. Can a gesture be an iconic representation if the referent of the icon does not appear? As long as the referent is present in the imagination (and keep in mind that only imagined aspects of the airplane appear in this event), the gesture can be coupled to someone's predictions in a way that makes it not only a representation but an iconic one. The instructor's gesture also has a relationship of mutual elaboration to the concurrently produced spoken word "Because." This relation is semantically complementary (metonymic), because the gesture represents an element of the procedure (trimming) that defines the feature (realism) that is the basis of the difference in the comparison schema. It soon became clear that trimming was not the aspect of the comparison

schema that PM went on to elaborate, and the instructor quickly abandoned the trimming gesture.

This gestural mismatch may have happened for one of two reasons. First, PM's projection of a reason for Boeing realism could have been illustrated with either a feature of the Boeing technique or by a feature of PM's company's technique. In choosing to model a feature of the Boeing technique, the instructor may have simply mistaken which continuation PM was projecting. However, the situation could be even more interactive. A second reason for the mismatch is that PM may have also been projecting a feature of the Boeing technique, but once this had been created by the instructor in the collaborative construction process, PM was free to provide the other meaningful completion. This interpretation also might rely on something like the Grice's (1975) maxim of quantity. Because the instructor had already illustrated the distinctive feature of the Boeing way, PM could increase the informativeness of his contribution by describing the distinctive feature of his company's procedure: "We always manage to keep our trim forward, you know, out of habit." PM can refer to this as keeping the trim "forward" because the stabilizer trim indicator is mounted on a horizontal surface at both sides of the center console. On that indicator, airplane nose-down trim is forward and nose-up trim is aft (Figure 6.2.6).

**Figure 6.2.6. Stabilizer Trim Indicator**
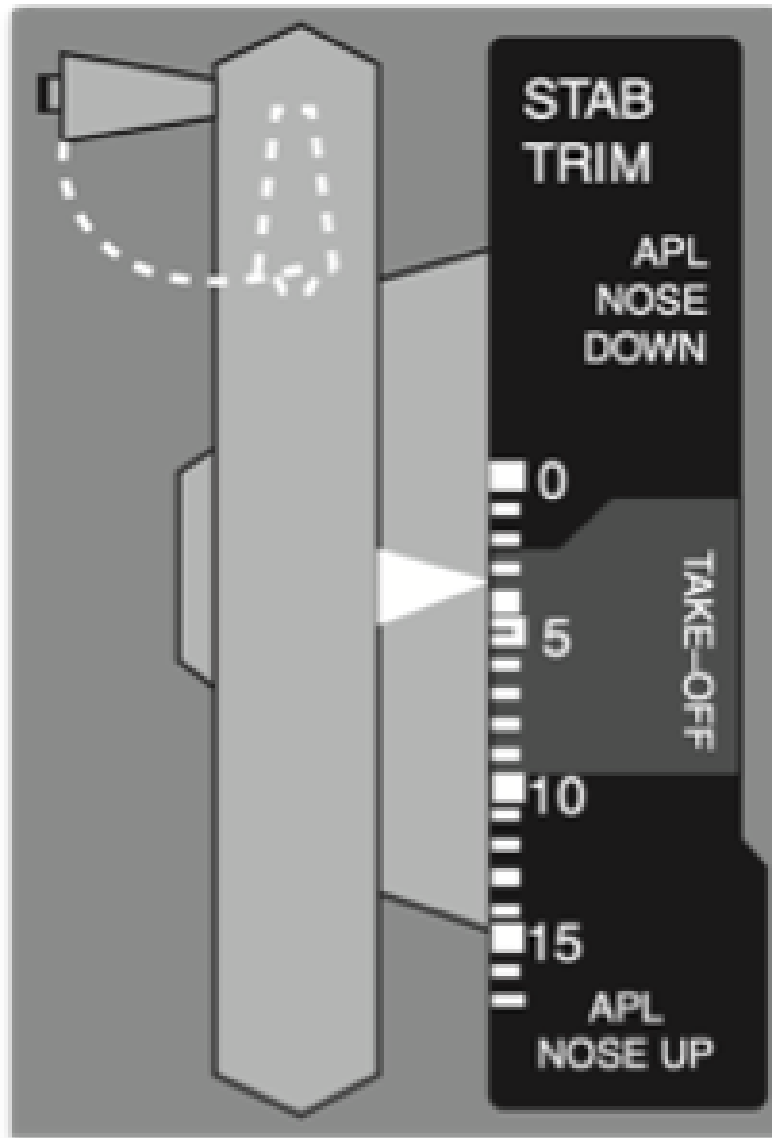
Notice that when talking to the PM, who would occupy the right seat in the simulator, the instructor modeled the trim action using his right thumb. The yoke-mounted trim switch is on the outboard horn of each control yoke. Thus, for a pilot in the right-hand seat (co-pilot's seat), the trim switch will be under the right thumb. Later in the same discussion, the instructor modeled pushing the thrust levers up with his right hand. This

gave his gesture an implicit body location in the left seat (PF's seat), and his subsequent gestural reference to trim was made with the left thumb. This coherence of the handedness of gesture with the seat occupied in the imagined flight deck indicates that the imagination of component actions, such as thrust changes and trim adjustments, involves the whole situation of the body in the flight deck, not just imagining the control that is to be manipulated.

## 6.2.4. You Have to Push

Conceptual object: To recover from stall attitude, push the yoke forward to cause nose-down pitch attitude. This was constructed from the PF character viewpoint.

The instructor created a role-playing narrative in which he modeled an inattentive pilot trimming into a stall. As the instructor finished his narrative, he continued to model the application of nose-up trim. PM began the following utterance speaking over the end of the instructor's narrative.

**Figure 6.2.7. It's really difficult**
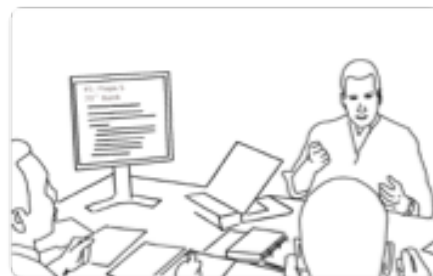


This complex example integrates seven gestures and five spoken utterances. All of the spoken elements and three of the gestures refer to the core conceptual object being constructed. Of the other four gestures, one refers to a previously developed conceptual object, one solicits agreement from another speaker, and two provide assessments of another speaker's conceptual project. All three pilots participated in the construction of this conceptual object.

Let's look first at the three push gestures produced by PM. While saying "It is very really difficult," PM modeled pushing the yoke forward. PM repeated the yoke-pushing gesture while saying "nose down." Finally, he said, "because you have to push," accompanied by a third pushing gesture. Each of the pushing gestures modeled pushing the yoke forward, and all are semantically congruent with the spoken words "to push" that occur at the end of the utterance. PM thus produced two anticipatory gestures followed by a third one that was produced simultaneously with the talk it elaborated. There are three content nodes represented in the speech stream. Each bears a different semantic relation to the conceptual content of the push gestures. The pilot action required to accomplish the recovery is represented by the spoken fragment, "you have to push." This spoken element bears a congruent relation to the push gestures. Pushing the yoke forward causes a nose-down pitch attitude represented in the spoken fragment "nose down." This spoken element bears a complementary (metonymic) relation to the push gestures. The recovery itself is represented by three spoken fragments, the first two produced by PM and the third produced by PF: "It's very really difficult," "back to normal," "To ah::: to recover from." These spoken elements bear a semantically complementary (synecdoche) relation to the push gestures.

At the beginning of this case, while the instructor was producing an iconic gesture as a follow-up to his previous narrative, PM changed the subject. The instructor stopped his trimming gesture after PM said "difficult" and modeled pushing the yoke. At this point, the instructor seemed to recognize the topic shift. The unexpected topic change created a relation of incongruence between the trimming gesture and the concurrent speech. This incongruity was not without meaning, however. Because the alignment of conceptual projects is an indication of membership in a shared community of practice, the divergence of co-imagined worlds produces a sense of exclusion. This was the second time that the instructor had anticipated a projection of PM's utterances that was not consummated (the first happened in case 6.2.3). The conceptual projects of the instructor and PM seemed less well aligned than those of the instructor and PF. This sort of interaction pattern could lead the instructor to feel that PM is not fully cooperative, even though he might be hard pressed to articulate an explanation for that feeling.

## 6.2.5. Underslung Engines

Conceptual object: Airplane dynamics; airplanes with engines mounted under the wings tend to pitch up when thrust increases. The instructor's gestures were constructed from a

character viewpoint, taking the speaker's body to be the airplane. PF's gestures were constructed from an observer viewpoint above and behind the wings, facing forward.

As with the previous case, this one is so complex that a full analysis is not possible here. In this case, all but one of the spoken elements and all of the gestures participate in the construction of the conceptual object. We can simplify the discussion somewhat by noting that the conceptual object has two principal parts: the location of the engines under the wing, and the pitch-up moment created by increasing thrust on engines that are so located. The instructor constructs the engine location by himself. The resulting pitch-up moment is collaboratively constructed by PF and the instructor.

**Figure. 6.2.8. These engines**

I: once those engines,
[*cupped hands at side below shoulders*]

they are under slung engines,
[*two beats with cupped hands at side below shoulders*]
right?

The engines, these.
[*arms extended out above previous engine gesture location*]

PF: (unintelligible)
[*fingertips of both hands rotate up quickly*]

"Underslung" describes a relationship between the engine and the wing. To create the relationship, the instructor used his body to enact the key parts of the airplane. He began by cupping his hands and holding them below his armpits while saying, "Once those engines."

It is possible that the gesture is iconic for the instructor before it becomes so for the students. The gesture becomes an iconic representation only when it is imagined as engines under a wing. The process is helped along by the gesture being produced concurrently with the words "those engines." At that point, the gesture can be seen as iconic of the engines and is semantically congruent with the words. The gesture was idiosyncratic and would have been quite ambiguous if taken in isolation. The words and gesture mutually elaborated each other. The words resolved the referent of the gesture (hands are engines), and the gesture contributed positioning information (the two engines are located in an imagined space here) that was not present in the words.

With the engines now located in an imaginary body-constructed space, the instructor elaborated on their location, simultaneously emphasizing the cupped-hand gestures with

Hutchins, E. (2025). A New Cognitive Ethnography. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.15828214
2025/4 - paris-ias-ideas - Article No.4. Freely available at https://paris.pias.science/article/a-new-cognitive-ethnography - ISSN 2826-2832/© 2025 Hutchins E.
This is an open access article published under the Creative Commons Attribution-NonCommercial 4.0 International Public License (CC BY-NC 4.0)

two vertical beats while saying, "they are underslung engines." This beat on the gesture and spoken fragment have a concurrent complementary relation; the gesture anchored the engines in a space, and the words implied something else (a wing) that had not yet been explicitly represented. The instructor then extended his arms out to the sides of his body, giving explicit representation to the previously implied wing, and said, "the engines, these." This gesture was positioned in space above the previously depicted location of the engines. Even though the space implied by the previous gestures and talk was now completely invisible and imaginary, it endured as a resource that could be exploited by subsequent meaning-making activities. Of course, the pilots know very well that the engines are located under the wings of the airplane. Once the perspective is established, the subsequent gestures and utterances highlight particular features of a familiar airplane configuration that has already been imagined. These words and gesture had a complementary semantic relation (gesture depicted the wing while speech referred to the engine) and were temporally concurrent. At this point, the construction of the location of the engines with respect to the wing was complete.

The fact that the space that was constructed by earlier actions could later give meaning to new gestures demonstrates that a discussion of pair-wise relations is fundamentally incomplete. We have picked out what appear to us to be the most significant relations, but our description remains partial because all of the elements of this complex semiotic field have important semantic and temporal relations to one another. The incremental construction of the airplane in a series of iconic gestures provides a nice example of the way that elements of discourse highlight selected portions of a rich, and never fully described, underlying model. We also see here not only anticipation of elements that are expected to appear in the future of the interaction, but retention of imagined meaning to provide context for subsequent interpretation.

**Figure. 6.2.9. Tend to sling this airplane up**



I: So, it's gonna.
[bends forward bringing wing gesture down]

PF: tend to, yeah
[flicks fingertips up again]

[I: entire body and arms come up]

I: hhh.
It's gonna tend to (0.3)
[bends at waist and lowers arms]
sling this airplane up.
[entire body and arms come up again]

In the context of the earlier discussion of the need to apply maximum thrust, the instructor's multimodal construction of the location of the under-wing location of the engines projected the airplane pitching nose-up in response to a rapid increase in thrust. PF used his two hands to model the rotation in the pitch axis caused by the increasing thrust on the two engines. His enactment was quite specific, showing the two engines and the torque that they would apply to the wings of the airplane when thrust was increased. Simultaneously, he said something that we have not been able to reconstruct. PF's gesture may have had congruent semantic relations with two spoken elements, one produced concurrently by PF himself and the other anticipated in the speech of the instructor. The instructor continued to develop his narrative, saying "So, it's gonna" while bending at the waist with his arms still extended to his sides. PF seemed to recognize this as preparation for a full-body stroke. A moment later, as the instructor swept his body and arms upward, PF flicked his fingers up again and said, "tend to, yeah." This gesture by PF is semantically congruent with an anticipated, but not yet produced, description by the instructor of the airplane pitching up. This gesture is temporally concurrent and semantically complementary (metonym) with PF's own words "tend to." PF performed this gesture in synchrony with the instructor's first full-body upward stroke. Thus, in addition to relations with spoken elements that were produced before, concurrently with, and after the gesture, PF's gesture also has a temporally concurrent and semantically congruent relation to the gesture produced by the instructor. Both gestures provided an iconic representation of the pitch-up event, but they were rendered from slightly different viewpoints. The instructor's viewpoint is as his body is the airplane itself, while PF imagines the pitch-up event from a viewpoint above and behind the airplane. It could be argued that gestures that have the same referent but are rendered from different actor viewpoints should be regarded as semantically complementary rather than congruent. At this time, we do not have a strong view on the matter. Simply posing the question highlights the possibility that semantic congruence is a continuous rather than a discrete function.

PF's utterance fragment "tend to, yeah" has a temporally concurrent and semantically complementary (metonymic: cause and effect) relation to the instructor's first full-body gesture. The instructor's second sweeping full-body gesture was produced concurrently with his own, now eagerly anticipated, verbal description of the pitch-up event, "sling this airplane up."

## 6.2.6 Discussion

When multiple authors speak and gesture together, the relationships of mutual elaboration proliferate. The extent to which participants become conscious of this wealth of meaning is currently unknown. We suspect, however, that the impression of complexity created by examining the relations among semiotic resources, one relation at a time, is somewhat misleading. From the participants' point of view, a single conceptual object emerges, and the many relations among the elements from which the object emerges fit naturally into the familiar structure of the conceptual object.

The participants are engaged simultaneously in two kinds of projects: They are enacting conceptual objects of interest (what they are talking about), and they are conducting a social interaction. Even though these projects are analytically separable, in action, they are woven into the same fabric. This was evident in case 6.2.4, where three of seven gestures modeled conceptual content, whereas the other four gestures accomplished speaker positioning in the interaction.

Surely pilots can imagine their work without speaking or gesturing. However, when they speak and gesture, the process of imagination becomes observable by and available to others. This is important for the participants because it allows them to collaboratively construct conceptual projects. It is critical for us as analysts because it enables us to record and analyze the process of conceptualization.

Gesture, talk, printed words, and material objects all have different representational affordances. Both talk and gesture are parts of the MGC process. Talk and gesture as actions involving multiple modes (M) in each participant are generated (G) as imagination and expectation by the models held by whoever produces them, AND are generated and imagined by those who witness the talk and gesture as part of the process of discovering their meanings. These actions feed back continuously (C) to all the participants' experiences of the models. Talking and gesturing about such things is a sort of exploration of the model for the speaker and a guided tour of the model for the recipients. In these vignettes, all of the participants are simultaneously explorers, guides, and guided.

In addition to modeling specific actions, many of the observed gestures presupposed specific flight deck roles, the seat occupied while performing the imagined action, and the fine details of the bodily motions of the pilot. Such details are rarely represented

linguistically in our data. The coherence of gestural enactments indicates that the imagination of component actions involves the whole situation of the body in the flight deck, not simply the imagination of the control that is to be manipulated. The richness and specificity of the pilot's shared knowledge of the flight deck environment are evident in the rapid shifts in viewpoint implied by the gesture sequences. Pilots transition seamlessly from character viewpoint to observer viewpoint, and among multiple vantage points as observers.

One way to bring relations of mutual elaboration into focus is to notice what does *not* appear in talk. For example, the control yoke, the trim switch, and the thrust levers play central roles in the interaction, yet these controls were never mentioned in the verbal utterances produced by the instructor and students. The controls are brought forth as implied elements in an imagined world of culturally meaningful action. The words "you have to push" could apply to many controls in the flight deck. That these words describe an action taken on the control yoke is established by their relation of mutual elaboration with particular gestures. And also, importantly, by the role the yoke plays in the conceptual project. Even without gestures, pilots know that the yoke is the thing that must be pushed in order to recover from an approach to a stall.

Gestures may enter into relations of mutual elaboration with many other semiotic resources in the activity system: written materials, objects, bodies, talk, and even other gestures. Gestures are complex movements. Just as we saw in the navigation team's choice of landmarks, which aspects of movement are taken to be relevant in the current moment of discourse depends on how the gesture is imagined as an element of the underlying model. For example, early in the lesson, the instructor read the words "retract speedbrake." The words say nothing about how the retraction of the speedbrake is accomplished. The speedbrakes are panels on the wings. Where is the activating control? How is the control operated? The instructor held his right fist upright in front of his body at elbow level. As he moved his hand forward, he rotated his wrist and tipped his fist down slightly. If this gesture were to occur alone, its meaning would probably be misunderstood. Viewed without sound, the gesture could easily be seen as modeling a pilot in the right seat pushing the right horn of the control yoke forward. But the gesture produced concurrently with the words "retract speedbrake" in this context brings forth an unambiguous whole. A pilot seated in the left seat of the flight deck uses his right hand to grasp the raised speedbrake handle and push it a few inches forward and down. Details of the motion that did not seem important when viewed without sound now jump out. The speedbrake handle pivots around a hinge at its base,

and this detail is shown in the gesture as the slight rotation of the wrist. Furthermore, details of the gesture that should be ignored fade away. In the airplane, the speedbrake handle is adjacent to the right thigh of the pilot in the left seat. A gesture that perfectly modeled speedbrake retraction would be performed below the pilot's waist. In the classroom, however, the surface of the table intervenes in the instructor's local space, preventing him from lowering his hand further. In mutual elaboration with the talk, the height of the gesture is disregarded. This is a reminder that even seemingly simple gestures may be extremely complex. What is meaningful and what is not, what should be attended to and what should be disregarded as noise, depends on how the gesture is construed. And the level of detail that can be achieved in the construal depends on the depth of knowledge that the participants have about the domain of discourse. It is not just the words and gesture that mutually elaborate each other. The words and the gesture are generated and imagined as elements in a dynamic constellation of activation in an MGC system that spans brain, body, and world. They are the visible parts of a complex imagining of a meaningful action in a known world. Talk and gesture are the tip of an MGC iceberg.

In the domain of professional pilot training, the participants use gesture to represent activities, objects, and events with respect to which all of the participants have thousands of hours of experience. Extensive embodied experience results in rich representational potential. Representational potential is realized in the enactment of the concepts in word and deed. Some of the meaningful flight deck actions are performed so often and so distinctively that the gestures derived from the actions attain the status of conventions in the community. Conventional status depends on the specificity of the gesture and its relations to other forms in the ecosystem. For example, nothing else that is done on the flight deck looks like holding two thrust levers (with a characteristic hand shape) and pushing them forward in a vertical arc that models the arc of the thrust lever quadrant. This actor-viewpoint gesture contains elements of both path and manner. The specificity of the gesture also depends on the standardization of the flight deck. Thrust levers and throttles are nearly universal in transport aircraft. Considering that virtually every airline pilot experiences the thrust levers in the same way, and that the bodily motions associated with manipulating the thrust levers are distinctive, this motion has gained the status of an iconic convention in the pilot community. Control yokes are not as widely distributed as thrust levers (having been replaced by side sticks in Airbus airplanes), but are still present in many airplanes and are understood by all pilots. The presence, position, and operation of many other controls are more variable across the

world's airplane fleets, and so, although the manipulation of these controls can be meaningfully modeled in gesture by pilots in context, they do not stand as interpretable context-independent iconic representations. The "retract speedbrake" gesture produced by the instructor, for example, is not so widely used as to be considered a convention.

Talk, gesture, and other semiotic resources are elements of an ongoing dynamic process of imagining activity in the workplace that is distributed across the pilots. Phenomena in the world are highlighted by, and acquire meaning from, gestures enacted in coordination with them. Simultaneously, gestures acquire meaning from the elements of the physical world with which they are coordinated (Goodwin, 1994, 2007). Of course, environmentally coupled gesture is pervasive when pilots work together on a flight deck, as we saw in the analysis of the diagnosis of the fuel leak. (See also Hutchins et al., 2009). Many of the gestures we observe in the pre-simulator briefing mutually elaborate physical elements of the briefing setting. But what of the gestures that refer to the absent flight deck? The fact that pilots have so much experience in this setting changes the dynamics of these processes. Once it has been invoked in speech or gesture, the entire flight deck becomes available (in imagination) as an environment to which subsequent gestures can be coupled. The same processes that are at work in meaning making with environmentally coupled gestures are at work here, except that these gestures both bring forth the imagined environment and are coupled to elements of that imagined environment. As we saw in the case of retracting the speedbrake, a gesture can selectively highlight elements of an imagined environment, while the imagined environment simultaneously draws attention to and gives meaning to subtle details of the gesture.

Gesture provides evidence that imagination can run ahead of talk (Schegloff, 1984). In case 6.2.4 "You have to push", PM made three yoke-pushing gestures but did not verbally describe the push action until the third gesture, seven seconds after the first push gesture was produced. The first two push gestures anticipated the semantically congruent spoken words. They were produced concurrently with semantically complementary elements of a verbal preamble that contextualized the pilot's stance with respect to the recovery maneuver (it's difficult) and with respect to the effect of the push (nose-down attitude). The third push gesture was produced concurrently with the words "to push." One consequence of repeating the gesture is that it kept the main point active while the verbal preamble was delivered. That is, the pilot was imagining the push action seven seconds before he got around to describing it verbally. The syntactic constraints of spoken language impose sequential order on the articulation of conceptual

elements. Gesture that anticipates one's own talk is a constituent of this pre-articulatory imagination. Before doing this analysis, we would have guessed that the most likely timing relation for collaboratively constructed multimodal utterances would be the production of words followed by semantically congruent gestures. Gestural follow-on assumes that the listener inhabits a conceptual world that is constructed *in response to* what the speaker has already said. Sometimes, however, gestures in collaboratively constructed multimodal utterances occur concurrently with the words they elaborate. In case 6.2.5 "Underslung engines," PF and the instructor executed perfectly synchronized, but morphologically distinct, enactments of a sudden pitch-up attitude. PF's gestures were performed in anticipation of the instructor's subsequent metaphorical description of the pitch-up event. The cross-speaker production of such multimodal elements in precise temporal and conceptual alignment requires joint participation in the embodied experience of this key conceptual element. It is further evidence that the participants jointly inhabit the world they imagine in interaction.

Simultaneity of cross-speaker gesture and talk in collaboratively constructed utterances is evidence of a shared activity and aligned expectations. This relation indicates that the speakers inhabit a shared conceptual world that is elaborated in parallel. As in single-speaker utterances, gestures in collaboratively constructed multimodal utterances often precede the spoken elements to which they bear semantic relations. The production of utterances in which a one speaker's gestures anticipate the conceptual projections of another speaker's words provides strong evidence that the participants inhabit a shared imagined world.

The details of such imagined worlds emerge incrementally as the semiotic resources of the setting are marshaled in interaction. When PM said (in case 6.2.2 "Back pressure only"), "Our procedure just trim out at flap five speed," he evoked an imagined world of a pilot preparing for the maneuver. The word "just" signals the absence of the further trimming below the flaps-5 speed that the previous discussion led one to expect. The instructor's next actor-viewpoint gesture showed that he had entered the imagined role of pilot created by PM. The instructor filled the projected conceptual hole by enacting the next part of the maneuver. He said, "and then" while modeling pulling back on the yoke.

Not all conversational projections are consummated. In case 6.2.3 "You have to push", PM introduced a conceptual scheme (a comparison) that could be developed in either of two ways. The instructor produced a gesture that committed to one projection of what

PM had said. PM went on to articulate the other projection. We have no evidence concerning PM's original intentions, but his actions do suggest that in the dynamic process of co-authorship of ideas, participants make choices in real time based on the shifting direction of the development of the conceptual object. A speaker may change the structure of a sentence while it is being produced. Furthermore, the choices can have social consequences, seeming to support or undermine the projects of the other person.

The occurrence of collaboratively constructed multimodal utterances demonstrates that the pilots jointly develop conceptual objects. The properties of this ecosystem create particular cognitive roles for gestures. In the cognitive ecology of flight training, some gestures link models among the pilots. Other gestures seem to be pre-enactments of actions that will be taken later when the pilots are in the flight simulator. When an action has more than one function, it may be that many functions are served simultaneously.

The pilots' bodies are a key resource in the process of conceptualizing their world and the actions they take in it. Conceptualization is not only multimodal but may also be a collaborative project; the modalities that are in coordination may be located in more than one participant. The range of possible relations of mutual elaboration among semiotic resources is extremely rich in collaboratively constructed multimodal utterances produced by experts engaged in consequential activity in a culturally constructed setting.

## 6.3. The Flight Crew Cognitive System

[12] As I noted in the background section, developments in the technologies of data collection and analysis change the phenomena we see, and that in turn provokes modifications to our theories. Over several years, the lab I shared with Jim Hollan at UCSD explored the creation of a digital ethnographer's workbench. To support the study of activity as an integrated system, our lab created a tool called ChronoViz. ChronoViz was the doctoral dissertation project of Adam Fouse, supervised by Jim Hollan and developed under a grant from NSF and a research contract with Boeing. For more details on ChronViz and its use in studying ongoing activity, see Fouse et al. (2011).

ChronoViz supports visualization and analysis of multiple sources of time-coded data, including multiple sources of high-definition video, simulation data, transcript data, paper notes, and eye gaze data. Each data source can be independently aligned with the

rest of the data and then used for navigation of the data set as a whole. Researchers using ChronoViz have rich interactive capabilities for exploring data sets, making annotations about observed activity, filtering and arranging annotations, and performing computational analysis of the recorded data.

In this section, I will present an analysis of just one event recorded in a Boeing 787 Dreamliner flight simulator. Both pilots wore eye-tracking glasses. Integrating dual eye-tracking data with the audio and video recordings of both pilots in ChronoViz permits us to describe the allocation of visual attention by the flight crew as a system.

**Figure 6.3.1. Pilot wearing eye-tracking glasses in Boeing 787 Flight Deck**



**Photo by the author**

Ensuring that an airplane follows its planned flight path is the formal responsibility of the PF. In an automated airplane, this is accomplished primarily via visual attention. However, though this is not one of PM's official duties, the eye gaze data show that the PM also monitored the flight path by directing his gaze to the altitude, vertical speed, and airspeed indicators on his primary flight display (PFD). Similarly, modifying the flight route while in flight is the formal responsibility of the PM. However, we observed that the PF also contributed visual and verbal attention to this task. Thus, each pilot went beyond his formal responsibilities by monitoring the actions of the other pilot.

While still more than 30 minutes from their destination, the crew programmed the flight management computer for the arrival procedure. This requires the crew to retrieve information from a navigation chart and enter the information into the airplane's flight management computer system. PM held the printed arrival chart in his hands and said, "okay, so after Helens, then we're gonna go to Battle Ground." This is a spoken representation of one leg of the arrival from a waypoint called HELNS to a waypoint called Battle Ground. While producing this utterance, PM's eyes first jumped around the chart, making several brief fixations before fixating on the depiction of the HELNS waypoint. He then fixated just below the depiction of the HELNS waypoint (along the path of flight) for about half a second, then further down the route of flight for another half second. This was followed by a saccade to the information box for the Battle Ground waypoint, where he fixated for almost a full second.

This short vignette allows us to see how PM goes about seeing the static marks on the chart as a representation of a dynamic flight route. The pilot's eye gaze follows the route, which is an emergent property of the projection of an imagined trajectory onto the waypoints depicted on the chart. This simultaneously gives the chart meaning as a representation of the route of flight, places the waypoints in sequence, and enacts the anticipated route of flight in the motor activity of eye gaze.

Figure 6.3.2 shows this interaction. At this moment, the flight route has been represented in the flight deck system in three ways: (a) as graphics and text on the arrival chart, (b) in the pattern of the PM's eye gaze over the chart, and (c) in the words of the PM. PM's eye gaze and speech reveal an intrapersonal dynamic configuration of visual and verbal attention in the inspection of the chart and the construction of the utterance. The representation of the route in speech lags the enactment of the representation of the route in eye gaze. This lag is expected because the pilot must know the identities of the waypoints before he can verbalize them. What is not so

obvious is the way the depiction of the route on the chart, in concert with the cultural practices of chart reading in the professional pilot community, provides the resources for the pilot to enact the route in motor activity (eye gaze) by reading the chart in a particular way.

**Figure 6.3.2. Three parallel representations of flight route**



[29:55.06]
PM: MK so after Helens,

Then we're gonna go to

Battle Ground.
[29:59.16]
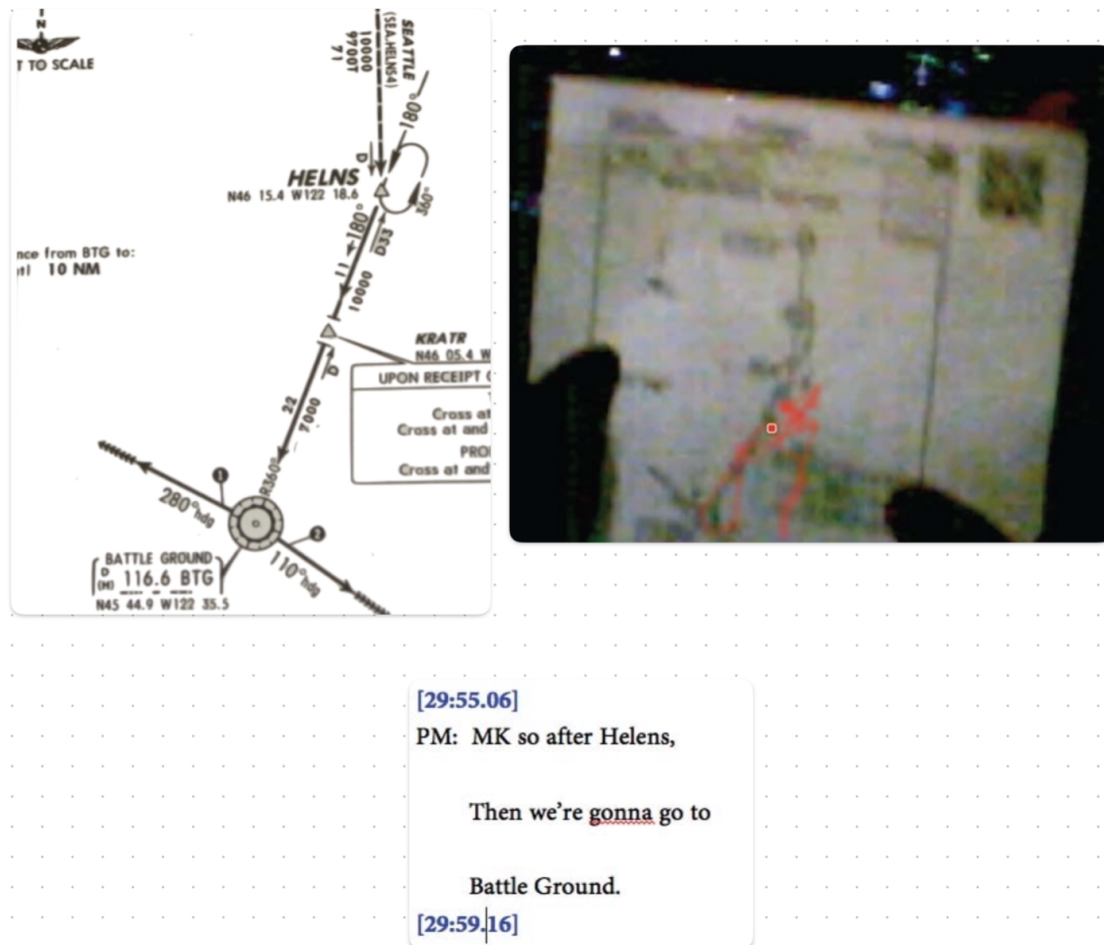
*Arrival Procedure Chart (upper left); PM's eye gaze (upper right); PM's verbal description (lower).*

Collecting eye-tracking data on both pilots simultaneously allows us to observe the allocation of visual attention by the flight crew system. While the PM was reading the chart as described above, the PF looked at the blank space on the computer interface where the name of the next waypoint after HELNS should be entered. The PM's spoken utterance thus coordinated the allocation of visual attention by the two pilots, each using a different cultural practice of reading, to two different representations of the flight

route, one on the chart and the other in the waypoint list on the computer interface. This coordinating effect is only possible because the language of sequential relations applies to both kinds of reading practices.

While the PM's visual attention followed a trajectory on the source of navigation information (the chart), the PF's visual attention was following a trajectory on the destination for that information (the computer interface). The PF's eye gaze anticipated the PM's next action in the activity, which was entering the identifier for the waypoint designated BTG into the list of waypoints that define the route. This complementary allocation of visual attention is evidence of the pilots' joint participation in and joint construction of a shared problem-solving activity.

As the PM then repositioned his body and hands to use the keypad to enter the waypoint identifier, PF shifted his eye gaze to the blank space on the waypoint list where the identifier for the next waypoint should be entered. At this point, the two pilots had produced not only congruent eye-gaze point of regard but congruent use of the practice of projecting a trajector on the display to read it as a sequential list.

At the moment, the PM read the chart while the PF read the waypoint list, the practices of the pilots were coordinated and complementary. When both looked at the same blank field on the waypoint list, their practices were congruent. This complementarity and congruence of practices determine important performance characteristics of the system.

The foregoing description provides details of the coordination of eye gaze for a brief interaction between the PF, PM, and airplane automation system. It is also possible to visualize the relationship of the eye gaze of the two pilots over longer spans of a flight. We use a gaze cross-recurrence plot to show this relationship (Coco & Dale, 2014). Figure 6.3.4 shows the gaze cross-recurrence plot for the flight from the performance of the pre-flight checklist up to and including the navigation problem solving described in the previous section.

To construct a cross-recurrence plot, we first define a set of Areas of Interest (AOI) in the region of the visual field where we have eye gaze data. The regions we chose are highlighted in colored boxes on the image of the B787 instrument panel, as shown in Figure 6.3.3.

**Figure 6.3.3. The seven areas of interest (AOIs) defined on the Boeing 787 Dreamliner instrument panel**

The eye gaze data for the two pilots is first mapped into the shared coordinate space of the instrument panel. To create the recurrence plot, we consider each fixation on an AOI made by the PF. We then find all instances when the PM fixated on the same AOI. For each such match, we plot a region in the recurrence plot located at the intersection of the temporal bounds of the PF fixation and the PM fixation and colored with the plotting color associated with the AOI, as shown in Figure 6.3.4. All colored regions indicate AOIs attended by both pilots. Colored regions on the diagonal indicate AOIs attended simultaneously by the two pilots. Regions to the upper right of the diagonal indicate AOIs attended by the PM before the PF. Regions to the lower left of the diagonal indicate AOIs attended by the PF before the PM.

**Figure 6.3.4. Cross-recurrence plot for 12 min of eye gaze in the Boeing 787 Dreamliner**
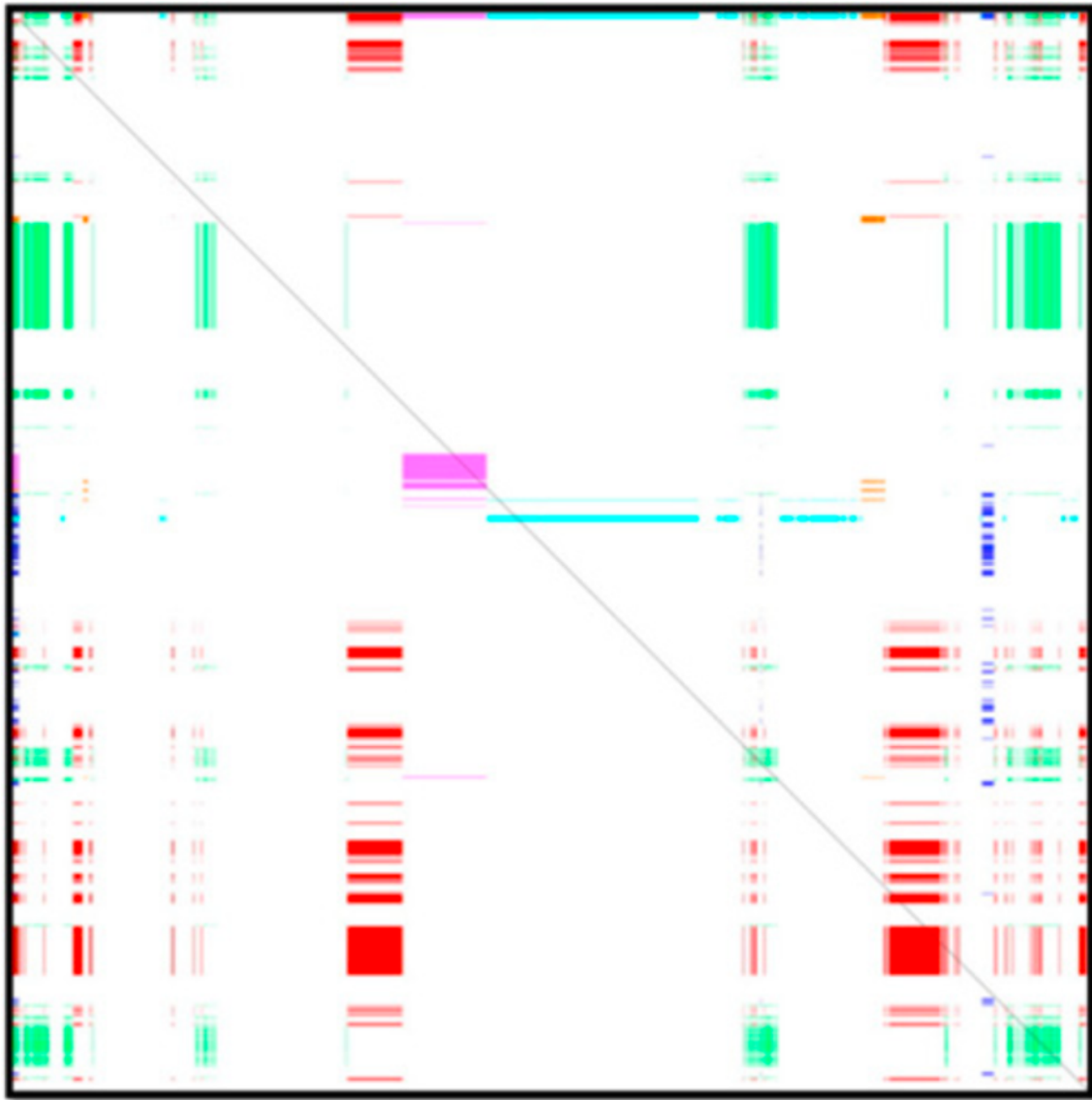
*The areas of interest (AOIs) attended by the pilot flying (PF) are shown in temporal sequence across the top of the plot from left to right, and the ones attended by the pilot monitoring (PM) are shown down the left side of the plot from top to bottom. AOI blocks on the diagonal were viewed simultaneously by PF and PM. Time increases from left to right, from top to bottom, and down the diagonal. Gaze fixations on AOIs are color-coded according to the scheme shown in Figure 6.3.3.*

The event begins with both pilots attending to the right side of the PM Navigation Display (ND). This region of the instrument panel is where the electronic checklist was displayed. These green boxes on the cross-recurrence plot indicate the visual attention allocated by both pilots to the joint performance of the preflight checklist. Following the diagonal down to the right, the large magenta box indicates the application of takeoff

thrust. The elongation of this magenta rectangle indicates that the PF begins attending to the engine instruments before the PM does. He also ceases attending to the engine instruments before the PM does, as is prescribed by the takeoff procedure. The next green region along the diagonal indicates the crew solving a navigation problem at that time. The PM's ND was configured in large map mode so that the region that, at other times, displayed an electronic checklist or the Control and Display Unit (CDU) was, at that time, part of the ND. The next red region indicates the crew again reasoning about the navigation problem jointly while looking at the PM's ND.

The final green region on the diagonal shows where the PM split his ND to display the CDU in the main instrument panel. On that display, the crew then jointly attempted, unsuccessfully, to locate and activate an approach to the destination airport. Both pilots abandoned this activity at the same moment, and the PM switched back to normal ND. This action is interesting because both pilots' eye gaze was entrained by the same display change. The entrainment gives rise to a square on the diagonal of the cross-recurrence plot.

This result suggests that the recurrence plot may be an excellent way to visualize the effects or effectiveness of cues that recruit joint visual attention. In some cases, the recruitment of visual attention of both pilots is desired or intended, as when a crew alerting system message is displayed, for example. In other cases, it is appropriate for the flight crew to allocate available visual attention to different locations. When the crew encounters a collision avoidance warning while flying in visual meteorological conditions, for example, it is not appropriate for both pilots to shift their gaze to the traffic display inside the flight deck. At least one pilot should be looking outside the flight deck, conforming to the "see-and-avoid" principle for visual flight conditions.

Modifying the flight route is a key automation management activity in all modern civil transport airplanes. Routes are specified to the flight management computer as strings of characters that specify sequences of geographic waypoints with associated altitudes and speeds. To build a route leg, pilots use the CDU and construct a sequence of waypoint designations and (if required) altitudes and speeds. One interesting aspect of this activity is that the pilots are required to reconcile two very different representations of the path of the airplane. Charts and navigation displays show the route graphically as points in space connected by route legs. The representation of the route on the CDU is a list of strings of characters, on which the only aspect of spatial layout that matters is that strings of characters higher on the page designate waypoints to be visited before those

displayed lower on the page. We did not expect to see a pilot enact the route of flight in the interaction of eye gaze with an arrival chart as the PM did, but once we saw that, it made sense. It was also not surprising to see a pilot's eye gaze enact the sequential spatial relations of route waypoint designations on the CDU route legs page as the PF did. After all, reading a list from top to bottom is a highly over-learned skill.

However, the coordination of these two kinds of eye gaze enactment of route legs in a joint navigation problem-solving activity was a surprise. The structure of problem solving and look-ahead is visible in the allocation of visual attention by the flight crew system, but *not in the gaze data for either pilot alone*. Consider PF's anticipation of PM's next action, for example. PF's eye gaze shift can be said to be anticipatory only in its temporal relation to PM's eye gaze shift. This finding means that the representation of the problem space in the flight crew system is not contained entirely inside the individual cognitive system constructed by either pilot.

The eye gaze cross-recurrence plot is generated by a computer program. It does require a specification of the AOIs, but once the AOIs have been defined, no frame-by-frame hand-coding of data is required. That is, the same analysis can be done with no additional work for similar data sets collected from any number of pairs of pilots. The gaze cross-recurrence plot provides a nice visualization of sequential aspects of the flight crew's visual attention system. When one pilot's eye gaze generally precedes or leads the other, it is visible as a weight of plotted regions to one side or the other of the diagonal. In fact, by computing the weight of plotted space on lines drawn parallel to the central diagonal on either side, one can quantify the extent of lead or lag between two subjects engaged in joint work. The gaze cross-recurrence plot also clearly shows the entrainment of eye gaze by salient display events. This property may make it a valuable tool in judging the effectiveness of crew alerting measures that are intended to draw the attention of both pilots.

The challenges of measuring, quantifying, and visualizing the performance of the flight crew *system* is at the heart of our program. The gaze cross-recurrence plot is an example of the measurement, quantification, and visualization of the properties of this system. This plot displays properties of the system of interaction rather than properties of either of the pilots in isolation. This is also a nice example of how advances in methodology raise new questions and challenges for the theory.

# 7. The Cultural Cognitive Ecosystem

In this paper, I developed a sketch of a theoretical cognitive system that is multimodal, generative, and is continuously coupled to an environment. This MGC framework incorporates lessons learned in my five-decade history of cognitive ethnographic research as well as ideas from a number of related fields on which I have drawn over the years. I used the draft MGC framework as a lens to scrutinize the fine-scale details of real-world activity. The framework was my license to reimagine the cognition taking place in those settings. I noted what could be seen through this lens and used what was seen to refine the MGC framework.

This entire enterprise was an experiment of sorts. It was an exploration of a familiar world seen from a novel perspective. I did not know in advance what would appear. Some things seen through the MGC lens appeared just as they were expected. Others were expected but seem to have better explanations under the MGC interpretation than they did previously. Still other phenomena appeared as surprises. I will not attempt an exhaustive recapitulation of the findings of the examination of activity. In the paragraphs below I describe a few selected phenomena that stand out in the MGC view.

## 7.1. A few findings that were enriched by the analysis

Let's begin with some phenomena that were expected, but not previously fully understood (by me, anyway) or explained well by classical views of cognition. These are things that look different through the lens of MGC. They are neither high-level nor low-level cognitive processes in the traditional sense. I consider the distinctions among levels to be a vestige of an earlier way of thinking about the mind. In the MGC view, this scale of levels of cognition is collapsed. It runs orthogonally to the Shallow/Deep and Concrete/Abstract dimensions of the MGC. For example, perception, being what happens when sensation meets expectation, involves the activation of neural circuits that are shallow and deep, concrete and abstract.

The MGC perspective on meaning-making differs from traditional accounts. The meaning of a flow of sensation is what is perceived, and perception is a product of the

sensation interacting with a prediction that arises from the state of the entire network below the sensorimotor surfaces. Perception is both a recognition process and a generative process. This generative aspect supports the projection of imagined structure onto the sensory traces caused by the presence of material phenomena in the world. This superposition of imagination onto perception (sensation + prediction) produces emergent phenomena that are not present in either the sensory traces or in the projected structure. This means that what is perceived may go well beyond the constraints provided by sensory experience. For example, the many scales in the navigation environment are rendered meaningful by the superposition of a trajector onto a spatial array. The G of MGC makes the link between material patterns and concepts. One SEES a pattern on sensory surfaces AS a concept by imagining the material pattern in the context of the activation of a representation of a concept deep in the network. The material pattern is recognized as an instance of a particular concept when the activation of that concept predicts the sensory trace of that material pattern.

The field of embodied interaction has long recognized that when multiple semiotic resources mutually elaborate one another (gesture, talk, action, external representations, etc.), the parts get their meanings from the whole. This observation runs counter to compositional models in which meaningful wholes are built up by combining independently meaningful parts or primitives. The idea that parts get their meanings from the wholes in which they participate actually seems impossible under many models of cognition. One appeal of predictive processing is that the mutual elaboration of semiotic resources can be easily imagined in a multimodal generative system that is continuously coupled to an actual or imagined world. The network settles to an interpretation that simultaneously satisfies the constraints imposed by all of the elements. Meanings can be vague, ill-formed, probabilistic, and perhaps even dynamically oscillating. Or they may be sharp and stable, brought into being by regions of activation that have clear and unchanging boundaries.

Just as perception and meaning making are products of both recognition processes and generative processes, so the handling of external representations involves both kinds of processes. Structures in the world may be "seen as" something other than themselves. We can see a number line as a sequence and see a line on the chart as a ship's track. The navigator sees a span of space between the tips of the dividers AS a distance in one moment and, a moment later, sees the same span as a speed. Navigators see a gesture over a chart as a potential line of position. Airline pilots see a finger

moving along a line painted on the fuel panel as a flow of fuel through a pipe. They see a thumb nodding above a fist as an adjustment of aircraft pitch trim.

Perceiving an external representation involves the sensory traces of the thing-in-itself and the patterns of network activation that recognize and predict the sensory traces of the thing-in-itself, PLUS the prediction, recognition, and imagination of the flow of sensation that would result were the represented thing also present. The internal patterns of network activation are those necessary to predict the sensation caused by the representation and by the thing represented. Representation is better thought of as resonance between internal and external processes than as maintaining an internal state that bears a resemblance to the thing represented.

It has long been recognized that concepts come to be manipulable as material patterns in external representations. The MGC perspective applied to the analysis of real-world activity shows how conceptual work can be accomplished through the sensorimotor engagement with external representations. The system that accomplishes this includes the world impinging on sensorimotor surfaces that provide an interface to a deep network. The network represents without resemblance the world on the other side of the senses. Predictive processing provides a mechanism to account for the catch phrase that 'perception is something we do, not something that happens to us.' As we have seen, thinking often happens in interaction with a material substrate. In the culturally organized material substrate, abstractions are made concrete and are thus made available to the senses and subject to manipulation. Cognitive processes that have traditionally been identified as 'high-level' can be carried out in the manipulation of physical objects.

Matter matters. Some external representations can be imagined in ways that allow offline processing that takes advantage of the structural properties of the once-external, now imagined, external representation. However, no one can imagine a chart or a compass rose with sufficient fidelity of do position fixing by mental imagery. For many tasks, the stable details of the external representation are necessary. The stability of external representations also provides stability to thought carried out while coupled to the representation. When thought concerns multiple variables that have many relations to one another, an external representation may be required to hold those relations without loss or transformation. Then, attention processes can sample from the external representation as needed to perform the task.

The probabilistic and context-dependent shifting of point of view shows up in many cultural practices. Consider the navigator using a chart. Beyond seeing the chart as a representation, other perspectives are at play. The body is present and positioned in some particular way with respect to the chart surface. Body space, constructed through proprioception, vestibular sensation, and muscle tension, maps onto the visual and haptic experience of chart space, which maps onto the space around the ship imagined from a bird's-eye perspective. The experience of this satellite perspective is produced by the navigator standing at the chart table looking down on the chart. These three spaces, body, chart, and world imagined from above, are simultaneously active and available with shifting emphases on them. The navigator also sees the world with a horizontal perspective when looking out the window of the navigation bridge and must reconcile that view with the world imagined from a bird's-eye (satellite) perspective on a chart. The view from the bridge involves the coordination of the body space, a bridge space (where is my body with respect to the layout of the bridge and the windows), and the world-outside-the-window space. Fluid movement among and reconciliation of these many spaces are apparent in the fine-scale details of the navigation team's activity.

Other computations are accomplished by shifting interpretations of physical states. The navigator superimposes a span between the tips of the dividers onto a numeric scale. He embeds this first in a network configuration in which the scale and the span represent distances. How far did the ship travel? Then, a fraction of a second later, he embeds that same scale and span in a network configuration in which they represent speed. How fast is the ship traveling?

The analysis of gesture can show how people fluidly shift point of view on imagined worlds. One pilot may configure his body such that his torso is the fuselage and his arms the wings of an airplane.Simultaneously, and in coordination with the first pilot, a second pilot sees the airplane from above and behind. His outstretched hands represent the pitch attitude of the airplane seen from that point of view. Meanwhile, a third pilot imagines himself to be seated in the cockpit, pushing the yoke forward while operating the trim switch with his thumb and looking down at the trim indicator. Goodwin (1994)said that all vision is perspectival. Imagination is perspectival as well.

People animate static external representations in a number of ways. Trajectories of attention applied to persistent configurations in material or imagined representations produce the linguistic phenomena known as fictive motion. We say that 'the road runs

down to the lake' even though the road does not move at all. The projection of an attention trajectory onto a set of numbers can construct a numerical scale. The superposition of a trajector of imagination on the ship's track was the basis of the inference that a new position plotted behind (as established by the imagined trajector) a previously position could not be correct.

Sometimes the trajectors that are projected onto material structures are enacted by body motion. In the airline flight deck, attention shifts embodied in eye gaze enacted the anticipated motion of an airplane on the approach plate and the sequential order on the RTE LEGS waypoint list. Eye gaze trajectories superimposed on different external representations revealed the coordinated activity of the two pilots. The dynamic process of fuel flow is enacted as a gesture superimposed on static painted lines on the fuel panel. Through motor resonance, this gave the other pilots a better representation of the situation being described by SO. The navigators created provisional lines of position by gesturing over the surface of the chart. This permitted them to imagine the angles of intersection of candidate triplets of LOPs.

Well-structured expectations can segment the flow of sensation. The problem of segmentation of the gesture stream was central in the discussion of navigators choosing landmarks to use for the next position fix. The choice criteria are constraints that generate predictions of gestural elements that may be depictions of lines of position. The sensory traces of meaningful landmark-like gestural strokes resonate deeply and are predicted. The sensory traces of motions without meaning (in the context of highlighting possible lines of position) dissipate. Prediction/expectation is a sort of filter that can isolate events in a continuous flow of experience and can distinguish relevant events from noise. There is some research on this topic, but more is needed (Benetti et al., 2023).

## 7.2. Some findings that came as surprises

Several of the plotting tools used on the navigation bridge have adjustable friction locks. These tools are used to record and hold temporary constraints. The degree of friction determines the robustness of the temporary memory, and since friction is adjustable, the navigator can control the duration of a memory through physical manipulation of the tool. Friction is thus a relevant resource for the control of thought. Managing friction requires fine motor control that is acquired through

experience with the tools. The skills of the navigator go well beyond knowing the procedures of navigation. It is likely that wherever a friction lock is present, embodied knowledge is at work. This insight came as a complete surprise.

Friction provides just one example of the complex interactions of the physical properties of the motor system with the dynamics of thought. When thinking happens in the manipulation of objects and tools in a culturally constructed environment for action, courses of action are trains of thought. The velocity profiles of movements could be determined by the physics of motor control, independent of the concepts being manipulated, or those velocity profiles might be in part determined by the conceptual organization of the task. It is likely that many actions are determined to varying degrees by both physical and conceptual constraints. Given the constraints of the physics of motion and the affordances of the tools being used, the dynamics of the motor system might affect what is easy to think and what is difficult to think. A lot remains to be investigated here. This line of research has implications for the design of environments for thinking. More must be known about this in order to judge what may be lost when the manipulation of analog tools is replaced by interactions with digital displays and controls.

All of the vignettes from the aviation domain involve motor resonance. Multiple actors act and gesture in the course of meaningful activities. We know that observation of those actions and gestures will activate the motor systems in the viewers. That was expected. The joint construction of a shared imagined world of action in the "approach to stall recovery" briefing vignette suggested that the phenomena go beyond motor resonance to a sort of cognitive resonance. There, we saw not just activation of the motor system as part of the conversational audience understanding gesture. The pilots were obviously exploring and contributing partial representations of a richly imagined world. Saeko Nomura and I had seen this shared imagined world in our earlier analysis of this activity (Hutchins & Nomura, 2011). Adding MGC analysis across the participants highlighted the cross-participant resonance. Each participant was maintaining a slightly different representation of the imagined world of action and events, and that representation was being tweaked, supported, augmented, added to, complemented, contradicted, and manipulated by the actions of other participants. As a coupled system, each pilot's imagination was manipulating and being manipulated by the imaginations of the other pilots.

In examining the interactions of navigators with their charts, or pilots with their cockpit displays I struggled with the question of whether and when the external representation was engaged as a thing-in-itself and when it was engaged as the thing it represents. I originally tried to identify contextual features that a person would use to decide between these two ways of engaging the representation. For example, I imagined the pilots wondering about a gesture, "Is this gesture about the fuel system in the airplane behind the cockpit, or is it about the fuel panel as a device in the cockpit?" Thinking more carefully about MGC, it became clear that there is no need for a discrete decision on this issue. Contextual features can bias the system in favor of one interpretation or the other. The degree of activation of the network regions that resonate with either interpretation is probabilistic and context-dependent. The data compelled me to imagine that multiple competing or simultaneously active interpretations are possible and even likely in the MGC system. The search for a clean distinction between these two modes of interpretation appears to be a case of mistaking a researcher's analytic device for a property of the system under study. That was a surprise.

I was surprised to find that several other distinctions that I had previously taken for granted dissolved under the MGC lens. The traditional distinctions between high-level and low-level processes, between action and cognition, between input and output, and between motor processes and thought processes all collapse or are seen to be probabilistic rather than dichotomous. Top-down and bottom-up passes of propagation of constraint are necessary for most computational instantiations of predictive processing, but they strike me as another example of mistaking a computational device for a property of the system under study.

There is one more surprise to mention. This is not a finding about the nature of cognition so much as a realization about my own thinking about cognition. For decades, I have had misgivings writing about representation. I was uncomfortable with what seemed to me to be cavalier assertions about internal representations. Working out MGC gradually led me to an understanding of internal representations not as things or stable patterns, but as functional abilities to imagine the world. That is, to generate or predict the sensory traces that would occur if the thing represented were present.Any such internal representation would only exist at those moments in which the imagination was being generated. Otherwise, the representation would be latent in the network, but not explicitly present. Predictive Processing and genAI demonstrate how this is a plausible story.

With respect to external representations, for some time I have argued that rather than asking "What is a representation?" we should be asking "When is a representation?" There is no list of necessary and sufficient features for the status of external representation. Some external pattern or process is only a representation when it is in a certain relationship to a person who is seeing that pattern or process as a representation of something that it is not. MGC gives me a better sense of the nature of the relationship between internal and external processes that make representation work. I now think of both internal and external representations as events rather than as entities with customary thing-i-ness. I could be wrong about representation, of course, but now that I have done the MGC analysis of many real-world activities, I'm surprised to find myself comfortable writing about representation.

## 7.3. Communities of minds in cultural worlds

Now, let's move our attention to a higher level of system integration, the cultural-cognitive ecosystem. It is a community of MGC systems engaged in culturally constituted activities and coupled to one another by virtue of the fact that, as they act in this world, they impose sensation on one another's sensory surfaces. This community is coupled, via action (including gesture and speech), to a material substrate that contains objects, artifacts, tools, machines, architecture, and inscriptions of all sorts: books, documents, and now digital devices.

As I pointed out in Cognition in the Wild (Hutchins, 1995), these interactions with social organization and material artifacts create system-level effects, including implementing Physical Symbol Systems.This is the secret of our success (and perhaps downfall) as a species. It gives us the power to control (and destroy) our environment. It's a good thing that socio-technical systems have different cognitive properties from individuals. If that were not the case, there would be no science and no advanced technology.

Documenting these emergent cognitive properties of groups and showing how they arise is an important job for cognitive ethnography. It is a necessary condition for a proper understanding of critical topics that range from the origins of language in our species to the cognitive properties of groups and institutions. It is an antidote to the attribution of the cognitive properties of the larger system to the individual. That miss-attribution (the fundamental attribution error baked into cognitive psychology) put cognitive science on

the wrong track for decades and continues to plague discourse about individual cognitive function.

Cultural practices are dynamic configurations of resources that mutually constrain one another and self-assemble through processes of constraint satisfaction. A practice is cultural if it involves the satisfaction of constraints that were created by social others. For humans, this covers almost everything a person does. This is how humans build their own cognition, and how they build on the accomplishments of others.

Participation in cultural practices induces dynamic cognitive processes that coordinate internal MGC activation with social and material resources. These processes, and the activities in which they arise, are generally cultural in the sense that they incorporate constraints that are the residua of the actions of social others. Culture is a community-level process that accumulates partial solutions to frequently encountered problems. Cultural practices are historically contingent, building opportunistically on what has been developed previously.

In this perspective, there is just one basic process, the generative predictive processing network operating continuously and simultaneously in every modality. This process participates in all instances of what have traditionally been known as high-level cognitive capacities. As I said in (Hutchins, 1980), decision making, memory, problem solving, reasoning, etc., are not different underlying processes; they are descriptions of task demands.

## 7.3.1. Dynamic Properties of the Cultural Cognitive Ecosystem

The cultural cognitive ecosystem is a dynamical system in which certain configurations of elements (what we know as stable cultural practices) emerge (self-assemble) preferentially. In this perspective, constraints on the assembly of these configurations exist in many places and interact with one another through a variety of mechanisms of constraint satisfaction. Some of these constraints arise from the complex interactions of neural systems as they encode beliefs, knowledge, categories, concepts, and so on. Other constraints reside in material stuff; documents, tools, furniture, towns, architecture, etc. as well as ephemeral patterns as those in sounds in music, warnings, and spoken language. Still other elements are emergent in social processes of collective intelligence,

the formation of a stadium wave, or residence patterns, or other patterns that emerge unintended from the individual decisions of many people.

The human cognitive system consists of a complex field of constraints internal and external. Practices both bring the constraints into contact with one another and provide the processes that satisfy the constraints. A constraint is not an object. It is an information process. The word 'constraint' describes the information regulation aspect of a process. A constraint is a process that rules some things out and rules other things in. More constraints satisfied means more states ruled in or ruled out, which implies greater organization, and that means greater predictability of experience.

Which practices assemble at any moment depends on the local structure of the ecosystem. That is, the occurrence of a practice depends on which constraints are available in local space and time.Experience, training, and the design of environments can all be seen as ways to bias the probability of the dynamic formation of particular practices. These things bias the assembly process by making certain constraints available.

Consider the chart table environment on the navigation bridge. It is a setting full of objects that afford certain interactions and not others. For each of these familiar objects, there are corresponding internal processes that anticipate and predict courses of interaction. The constraints operating at any moment can be internal, external, or both. This is my take on the problem of "recruitment of resources." Such recruitment is not scripted by the nervous system alone. Remember the plotter's discovery of the missing computational term. His "Aha! insight" emerged from the unintended superposition of motor anticipation on a visual field. Multimodal engagement of the world by an individual or by multi-person interaction are both ways to bring together constraints in novel ways.

The stability, resilience, or persistence of a practice depends on the network of relations to other practices within which it is embedded. This includes membership in a family of practices that share an underlying cognitive strategy. The generalization of action across contexts is facilitated by family resemblances among practices. There are different kinds of connections to be exploited when families of practices exist. Where there are many members in a large family of practices that share a cognitive process (for example, the superposition of a trajector on a spatial array), a person will have many opportunities to engage in the shared cognitive process, leading to facility with and refinement of the

skill. The generalization of action across contexts is facilitated when a single underlying skill underwrites differing practices. For example, search on a number sequence that appears in the pages of a book or streets in a town. These simple processes can be building blocks of cognitive assemblies that accomplish a very wide range of conceptual processes. Interpersonal coordination of practices, including communicative practices, is facilitated by the fact that families of practices exist.

The dynamics of practice formation and maintenance may include positive feedback loops such that the more prevalent a practice becomes, the more probable its formation. This sort of loop can control diversity, limiting the number of discrete practices. Like trails in a forest, cultural practices stabilize through a process of modulated positive feedback. The deposition of the ideal and material residua of practice may make the practice more likely in the future. This will increase the predictability of experience and is a means to control system entropy. These processes may also preclude the formation of other potentially useful practices.

Culture is learnable because the ecosystem of practices has structure. Possibilities for individual learning depend on the structure of the ecosystem, both because the local ecosystem determines the inventory of available things to be learned and because family resemblances among practices reduce the complexity of learning processes. Interlocking relations among practices may produce conditions of multiple determination or even over-determination of particular features. Learning in the ecosystem includes changes that are outside of individual persons, in the historical development of concepts and artifacts, for example. The richness of the ecosystem creates conditions of multiple determination that promote the reliable induction of internal elements in individuals. Generative systems, like other neural networks, learn more robustly when the outcomes or predictions are multiply determined.

## 7.3.2. Evolution in the cultural cognitive ecosystem

As is the case in all evolutionary systems, the development of new practices is constrained by the existing networks in the ecosystem. There are large-scale examples of this in the history of virtually every tool. For example, the practice of plotting positions using visual bearings depends on a circular degree scale, a direction frame of reference, a suite of tools, and the Mercator projection chart. A small-scale example is the three-minute rule. The three-minute rule is based on a regularity that is so robust it is

virtually impossible to NOT discover it. Once the convention of measuring small distances in yards arose (for reasons unrelated to this rule), whenever fixes are taken on three-minute intervals (also, originally, for reasons unrelated to this rule), the speed of the ship will equal the number of hundreds of yards travelled between fixes. An easily measured distance can be read as an otherwise difficult to compute speed.

A model of the evolution of cultural practices in a cultural cognitive ecosystem will be a much more complex model than the cultural evolution of traits or memes. Models of the evolution of traits or memes are computationally tractable. What would be entailed by taking cultural practices to be the central concept in a model of cultural evolution? We would have to consider the evolution of entire ecosystems rather than the evolution of individual cognitive species.

# 8. The other 90% of the human cognitive system

In the first decade of the 21st century, medical researchers recognized that many non-human organisms colonize the human body. Most of these colonists are bacteria. After a century of trying to promote health by killing germs, it is now recognized that many bodily functions rely on symbiotic relationships among human and non-human cells. In fact, a census of cells in and on the human body shows that about one cell in 10 is human. The other 90%, the so-called "microbiome," are non-human cells. Some researchers claim that a study of human health that ignores the microbiome is fundamentally incomplete. They predict that in the coming decades, the study of the interactions between the human body and its microbiome, the "other 90% of the human physiological system," will transform medicine (Gilbert et al., 2018; Turnbaugh et al., 2007).

An analogous situation exists in contemporary cognitive science. For most of its history, cognitive science has sought the human cognitive system inside the human brain. In an attempt to get a better look at what were thought to be the "real" cognitive processes, the contamination of the material and social worlds was scrubbed out of the experimental booth. We now know that such an approach is fundamentally incomplete. Many critical cognitive functions rely on interactions between the brain, body, and culturally organized world.

Finally, there is something that was not a surprise, per se, but I was nevertheless amazed by its strength. The analyses of the vignettes made the role of culture in the construction of human cognition more obvious than ever. This idea was not just strengthened by these analyses; I consider it fully established and undeniable. You and I and every person are, among many other things, MGC systems.Each of us is continuously in multimodal contact with a culturally constructed social and material environment. Internal generative processes are coupled to this world, continuously shaped by and shaping perceptions. People are not so much immersed in culture as they are woven into the fabric of culture. By virtue of co-causing the organization of a resulting perception, internal processes and external objects and events are linked to one another. Internal generative processes, being continuously shaped by culture, are themselves cultural processes. Imagine a coupled set of internal processes and external events as a single thread. These threads run through the people and their world. The threads constitute and are constituted by the cultural processes. When people interact, they create and grasp a shared thread. We are bound to one another by the threads of cultural process that weave us into the fabric of culture.

The analyses of the vignettes provide ample evidence that the MGC perspective provides new insights about culture and cognition. The findings of this experimental cognitive ethnography are not a different set of facts under the old way of thinking; they are the product of a new way of thinking. The paradigm shift I highlighted in the introduction will require an adjustment of the underlying assumptions of cognitive psychology and cognitive science. It will require rebuilding our understanding of culture, the mind, and cognition.

Much work remains to be done. In a paper examining how neuroscientists learn to interpret maps of brain activity, Morana Alač and I noted that studying action as cognition in real world settings should be "the natural province of cognitive anthropology" (Alač & Hutchins, 2004, p. 659). I hope this paper provides cognitive anthropologists and cognitive scientists in general with a starting point for a new cognitive ethnography.

# Bibliography

Agha, A. (2007). *Language and Social Relations*. Cambridge University Press.

Alač, M., & 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.

Barbasi, A. (2002). *Linked: the new science of networks*. Perseus Publishing.

Barsalou, L. (2010). Grounded Cognition: Past, Present, and Future. In *Topics in Cognitive Science* (pp. 716–724).

Bateson, G. (1972). *Steps to an Ecology of Mind*. University of Chicago Press.

Beer, R. (2008). The Dynamics of Brain-Body-World systems: a status report. In P. Calvo & A. Gomila (Eds.), *Handbook of Cognitive Science: An embodied approach* (pp. 99–120). Elsevier.

Benetti, S., Ferrari, A., & Pavani, F. (2023). Multimodal processing in face-to-face interactions: A bridging link between psycholinguistics and sensory neuroscience. *Front. Hum. Neurosci, 17*.

Bourdieu, P. (1977). *Outline of a Theory of Practice*. Cambridge University Press.

Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence, 47*, 139–159.

Calvo, P., & Gomila, T. (2008). *Handbook of cognitive science: An embodied approach*. Elsevier.

Campbell, M., & Cunnington, R. (2017). More than an Imitation Game: Top-down modulation of the Human Mirror System. *Neuroscience and Behavioral, Reviews(75*, 195–202.

Cangelosi, A., & Parisi, D. (2002). *Simulating the evolution of language*. Springer-Verlag.

Churchland, P., Ramachandran, V., & Sejnowski, T. (1994). A Critique of Pure Vision. In C. Koch & J. Davis (Eds.), *Large-scale neuronal theories of the brain* (pp. 23–60). MIT Press.

Clark, A. (1998). *Being There: Putting brain, body, and world together again*. MIT Press.

Clark, A., & Chalmers, D. (1998). *The Extended Mind* (pp. 7–19). Analysis.

Clark, A. (2001). *Mindware: An Introduction to the Philosophy of Cognitive Science*. Oxford University Press.

Clark, A. (2016). *Surfing Uncertainty: Prediction, action, and the embodied mind*. Oxford University Press.

Clark, A. (2023). *The Experience Machine: How our minds predict and shape reality*. Pantheon Books.

Coco, M., & Dale, R. (2014). Cross-recurrence quantification analysis of categorical and continuous time series: an R package. *Frontiers in Psychology*, *5*(510).

Cole, M. (1996). *Cultural Psychology: A once and future discipline*. Basic Books.

Daniels, H., Cole, M., & Wertsch, J. (2007). *The Cambridge companion to Vygotsky*. Cambridge University Press.

Dasen, P. (1972). Cross-Cultural Piagetian Research: A Summary. *Journal of Cross-Cultural Psychology*, *3*(1), 23–39.

Dreyfus, H. (1982). *Husserl: Intentionality and cognitive science*. MIT Press.

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

Dupuy, J. (2000). *The mechanization of the mind: On the origins of cognitive science*. Princeton University Press.

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

Engel, A., Stewart, O. G., & Paulo, E. (2010). *Directive Minds: How dynamics shapes cognition* (J., Ed.; pp. 219–244). MIT Press.

Fauconnier, G. (1997). *Mappings in Thought and Language*. Cambridge University Press.

Fauconnier, G., & Turner, M. (2002). *How We Think*. Basic Books.

Flash, T., & Berthoz, A. (2021). Space-time geometries for motion and perception in the brain and the arts. *Nature*.

Fouse, A., Weibel, N., Hutchins, E., & Hollan, J. (2011). ChronoViz: A system for supporting navigation of time-coded data. *Extended Abstracts of CHI 2011, SIGCHI Conference on Human Factors in Computing Systems*.

Friston, K. (2008). *Heirarchical Models in the Brain*. PLOS Computational Biology.

Gallese, V., & Lakoff, G. (2005). The Brain's Concepts: The role of the sensori-motor system in conceptual knowledge. *Cognitive Neuropsychology, 22*(3), 455–479.

Gentner, D., & Stevens, A. (1983). *Mental Models*. Lawrence Erlbaum Assoc.

Gibbs, R. (2006). *Embodiment and Cognitive Science*. Cambridge University Press.

Gibson, J. (1986). *The Ecological Approach to Visual Perception*. Lawrence Erlbaum Associates.

Gilbert, J., Blaser, M., Caporaso, J., Jansson, J., Lynch, S., & Knight, R. (2018). Current Understanding of the Human Microbiome. *Nat Med, 24*(4), 392–400.

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

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Harvard University Press.

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

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

Goodwin, C. (2000). Action and Embodiment within Situated Human Interaction. *Journal of Pragmatics*.

Goodwin, C. (2007). Environmentally Coupled Gestures. In S. Duncan, J. Cassell, & E. Levy (Eds.), *Gesture and the Dynamic Dimension of Language* (pp. 195–212). John Benjamins.

Goody J. (1977). *The Domestication of the Savage Mind*. Cambridge University Press.

Greeno, J., & Moore, J. (1993). Situativity and symbols: Response to Vera and Simon. *Cognitive Science*, *17*, 49–59.

Greeno, J. (1998). The situativity of knowing, learning, and research. *American Psychologist*, *53*(1), 5–26.

Grice, H. (1975). Logic and Conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics* (pp. 43–58). Academic Press.

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

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

Hazlehurst, B., & Hutchins, E. (1998). *The emergence of propositions from the coordination of talk and action in a shared world* (Vol. 13, pp. 373–424).

Heidegger, M. (1962). *Being and time*. Harper.

Holder, B., & Hutchins, E. (2001). What Pilots learn about autoflight while flying on the line. *Proceedings of the 11th International Symposium on Aviation Psychology*.

Hollan, J., Hutchins, E., & Weitzman, L. (1984). STEAMER: An interactive inspectable simulation-based training system. *AI Magazine*, *5*(2), 15–27.

Holland, D., & Quinn, N. (1987). *Cultural Models in Language and Thought*. Cambridge University Press.

Hurford, J., Studdert-Kennedy, M., & Knight, C. (1998). *Approaches to the evolution of language: Social and cognitive bases*. Cambridge University Press.

Hurley, S. (1998). *Consciousness in action*. Harvard University Press.

Hutchins, E. (1980). *Culture and Inference: A Trobriand Case Study*. Harvard University Press.

Hutchins, E., & McCandless, T. (1982). *Manboard: A graphic display program for training relative motion concepts* [Technical Report TN 82-10,]. Navy Personnel Research.

Hutchins, E., & Klausen, T. (1986). Distributed cognition in an airline cockpit. In Y. E. Middleton (Ed.), *Cognition and communication at work* (pp. 15–34). Cambridge University Press.

Hutchins, E. (1995). *Cognition in the Wild*. MIT Press.

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

Hutchins, E., & Hazlehurst, B. (1995). How to invent a lexicon: The development of shared symbols in interaction. In N. Gilbert & R. Conte (Eds.), *Artificial societies: The computer simulation of social life* (pp. 157–189). UCL Press.

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

Hutchins, E., & Holder, B. (2000). Conceptual models for understanding an encounter with a mountain wave. *Proceedings of HCI-Aero 2000*.

Hutchins, E., & Hazlehurst, B. (2002). Auto-organization and emergence of shared language structure. In A. Cangelosi & D. Parisi (Eds.), *Simulating the evolution of language* (pp. 279–305). Springer-Verlag.

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 N. Enfield & S. Levinson (Eds.), *Roots of human sociality: Culture, cognition and interaction* (pp. 375–398). Berg Publishers.

Hutchins, E., Nomura, S., & Holder, B. (2006). The ecology of language practices in worldwide airline flight deck operations. *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, 363–368.

Hutchins, E. (2007). Measuring change in pilots' conceptual understanding of autoflight. *Proc. of the 14th International Symposium on Aviation Psychology*, 281–286.

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

Hutchins, E., Middleton, C., & Newsome, W. (2009). Conceptualizing spatial relations in flight training. *Proceedings of the 15th International Symposium on Aviation Psychology*.

Hutchins, E. (2010). Cognitive Ecology. In *Topics in Cognitive Science* (pp. 705–715).

Hutchins, E. (2010). Enaction, imagination, and insight. In J. Stewart, O. Gappene, E. Paolo, & Enaction (Eds.), *Towards a New Paradigm in Cognitive Science* (pp. 425–450). MIT Press.

Hutchins, E., & Nomura, S. (2011). Collaborative Construction of Multimodal Utterances. In J. Streeck, C. Goodwin, & C. LeBaron (Eds.), *Embodied Interaction: Language and body in the material world* (pp. 29–43). Cambridge University Press.

Hutchins, E. (2012). Concepts in practice as sources of order. *Mind, Culture and Activity*, *19*(3), 314–323.

Hutchins, E., Weibel, N., Emmenegger, C., Fouse, A., & Holder, B. (2013). An integrative approach to understanding flight crew activity. *Journal of Cognitive Engineering and Decision Making, 353–376*.

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

Johnson, M. (1987). *The Body in the Mind: the bodily basis of meaning, imagination, and reason*. University of Chicago Press.

Kelso, S. (1995). *Dynamical Patterns*. MIT Press.

Kugler, P., & Turvey, M. (1987). *Information, Natural Law, and the Self-assembly of Rhythmic Movement*. Earlbaum.

Kutas, M., & Federmeier, K. (1998). Minding the body. *Psychophysiology*, *35*, 135–150.

Lakoff, G., & Nuñez, R. (2000). *Where Mathematics Comes From: How the embodied mind brings mathematics into being*. Basic Books.

Langacker, R. (1987). *Foundations of Cognitive Grammar* (Vol. 1). Stanford University Press.

Latour, B. (2005). *Reassembling the Social: An introduction to Actor-Network-Theory*. Oxford University Press.

Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life*. Cambridge University Press.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.

Lee, D. D. (1949). Being and value in a primitive culture. *Journal of Philosophy*, *48*, 401–415.

Malinowski, B. (1965). *Coral Gardens and their Magic: The language of magic and gardening* (Vol. 2). Indiana University Press.

Malone T., & Bernstein M. (2015). *Handbook of Collective Intelligence*. MIT Press.

Marchese, F. (2011). Exploring the Origins of Tables for Information Visualization. *15th International Conference on Information Visualisation*, 395–402.

Maturana, H., & Varela, F. (1987). *The Tree of Knowledge: Biological roots of human understanding*. New Science Library.

McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. University of Chicago Press.

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

Minsky M., & Papert S. (1988). *Perceptrons*. MIT Press.

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

Núñez, R., & Sweetser, E. (2010). With the Future Behind Them: Convergent Evidence From Aymara Language and Gesture in the Crosslinguistic Comparison of Spatial Construals of Time. *Cognitive Science*, *30*(3), 401–450.

Newell, A., & Simon, H. (1972). *Human Problem Solving*. Prentice Hall.

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

Nomura, S., Hutchins, E., & Holder, B. (2006). The Uses of Paper in Commercial Airline Flight Operations. *Proceedings of the 2006 ACM Conference on Computer Supported Cooperative Work*,

*CSCW 2006*, 249–258.

O'Regan, J., & Noë, A. (2002). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, *24*(5).

Palmer, E., Hutchins, E., Ritter, R., & Cleemput, I. (1993). Altitude deviations: breakdowns of an error tolerant system. *NASA Technical Memorandum*, *108788*.

Pea, R. (1996). Practices of distributed intelligence and designs for education. In G. Salomon (Ed.), *Distributed cognitions: Psychological and educational considerations* (pp. 47–87). Cambridge University Press.

Pfeifer, R., & Bongard, J. (2007). *How the body shapes the way we think: A new view of intelligence*. MIT Press.

Port, R., & Gelder, T. (1995). *Mind as Motion: Explorations in the dynamics of cognition*. MIT Press.

Rogoff, B. (2003). *The cultural nature of human development*. Oxford University Press.

Rowlands, M. (2006). *Body language: Representation in action*. MIT Press.

Rumelhart, D. (1975). *Notes on a schema for stories* (D. G. Collins, Representation, & Understanding, Eds.). Academic Press.

Rumelhart, D., & McClelland, J. (1986). *Parallel Distributed Processing*. MIT Press.

Rumelhart, D., Smolenski, P., McClelland, J., & Hinton, G. (1986). Schemata and sequential thought processes in PDP models. In J. McClelland & D. Rumelhart (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (Vol. 2). MIT Press.

Schegloff, E. (1984). On some gestures' relation to talk. In M. Atkinson & J. Heritage (Eds.), *Structures of Social Action* (pp. 266–296). Cambridge University Press.

Senft, G. (1986). *Kilivila: The Language of the Trobriand Islanders*. Mouton de Gruyter.

Simon, H., & Kaplan, C. (1989). *Foundations of Cognitive Science* (M. Posner, Ed.). MIT Press.

Smith, L. (2005). *Action alters shape perception*. Cognitive Science.

Spivey, M. (2007). *The Continuity of Mind*. Oxford University Press.

Sporns, O., & Zwi, D. (2004). The small world of the cerebral cortex. *Neuroinformatics*, *2*, 145–162.

Stewart, J., Gapenne, O., & Paolo, E. (2010). *Enaction: Toward a new Paradigm for Cognitive Science*. MIT Press.

Streeck, J. (1993). Gesture as Communication I: Its Coordination with Gaze and Speech. *Communication Monographs*, *60*(4), 275–299.

Streeck, J., Goodwin, C., & LeBaron, C. (2011). *Embodied Interaction: Language and body in the material world*. Cambridge University Press.

Suchman, L. (1987). *Plans and Situated Actions: The problem of human-machine communication*. Cambridge University Press.

Suleyman, M., & Bhaskar, M. (2023). *The Coming Wave: Technology, Power, and the Twenty-First Century's Greatest Dilemma*. Crown Books.

Sunstein, C. (2007). *Infotopia: How many minds produce knowledge*. Oxford University Press.

Surowiecki, J. (2004). *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations*. Doubleday.

Talmy, L. (1995). Fictive motion in language and perception. In P. Bloom, M. Peterson, L. Nadel, & M. Garrett (Eds.), *Language and Space* (pp. 307–384). MIT Press.

Thelen, E., & Smith, L. (1994). *A Dynamical Systems Approach to the Development of Cognition and Action*. MIT Press.

Theureau, J. (2015). *Le Cours D'Action: L'Enaction & L'Expérience*. Octares Editions.

Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the sciences of mind*. Harvard University Press.

Turnbaugh, P., Ley, R., Hamady, M., Fraser-Liggett, C., Knight, R., & Gordon, J. (2007). *The Human Microbiome Project. Nature*, *449*(18), 804–810.

Varela, F., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. MIT Press.

Vygotsky, L. (1978). *Mind in Society*. Harvard University Press.

Weibel, N., Fouse, A., Emmenegger, C., Kimmich, S., & Hutchins, E. (2012). Let's look at the cockpit: Exploring mobile eye-tracking for observational research on the flight deck. *Proc. ETRA 2012, ACM Symposium on Eye Tracking Research and Applications*.

Wertsch, J. (1985). *Vygotsky and the Social Formation of Mind*. Harvard University Press.

Wilson, S., Saygin, A., Sereno, M., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience,* 701–702.

# Footnotes

**1 :** Funding for this project was provided by the Social Science Research Council. My wife and I lived in Tukwaukwa Village on the island of Boyowa from June of 1975 until October of 1976. ↩

**2 :** My access to ships in the US Navy fleet was made possible by my position as a Personnel Research Psychologist at the Navy Personnel Research and Development Center in San Diego, California from 1980 through 1986. I worked in the Future Technologies Group under the direction of James Hollan.↩

**3 :** In 1988 Don Norman and I began a research project titled, "Distributed Cognition in Aviation." The project was funded by the Aerospace Human Factors Research Division of NASA under the Aviation Automation Safety program. I continued with NASA funding until 2001. Over the years I had research agreements including flight deck access on United Airlines, Delta Airlines, America West Airlines, US Airways, American Airlines, Aeroméxico, GOL (Brazil), All Nippon Airlines, Japan Airlines, KLM Royal Dutch Airlines, and Air New Zealand. I have had consulting agreements and research funding from both Boeing and Airbus. In corporate aviation I have had consulting arrangements with LOFT flight training in Carlsbad, California.↩

**4 :** This section draws on material developed for the article (Hutchins, 2010a).↩

**5 :** This is the same pseudonym I used for this ship in (Hutchins, 1995). ↩

**6 :** This is an original analysis done with the MGC framework in mind. It is more detailed and perhaps reveals more insights from the MGC framework than the other analyses in this paper.↩

**7 :** The analysis in this section is based on an analysis that appears in the early pages of (Hutchins, 2010b). The Santa Fe Institute's program on Robustness in Social Systems provided funding for the work reported here. Erica Jen served as grant monitor.↩

**8 :** The analysis in this section is based on (Hutchins, 2010b). 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 5.3.2 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 analysis.↩

**9 :** The analysis in this section is based on an analysis that appeared in (Hutchins, 2006). This work was funded by a grant from the Santa Fe Institute's program on Robustness in Natural and Social Systems. Alisa Durán transcribed the data and suggested many elements of the analysis presented here.↩

**10 :** The analysis in this section is based on (Hutchins & Palen, 1997).↩