I See What You Are Saying: Action as Cognition in fMRI Brain Mapping Practice*

MORANA ALAČ and EDWIN HUTCHINS**

ABSTRACT

In cognitive neuroscience, functional magnetic resonance imaging (fMRI) is used to produce images of brain functions. These images play a central role in the practice of neuroscience. In this paper we are interested in how these brain images become understandable and meaningful for scientists. In order to explore this problem we observe how scientists use such semiotic resources as gesture, language, and material structure present in the socially and culturally constituted environment. A micro-analysis of video records of scientists interacting with each other and with fMRI images reveals *action as cognition*, that is, actions that constitute thinking for the scientists.

Introduction

The conduct of scientific research involves many kinds of cognitive processes. Some of these are internal processes of the sort that have been the focus of cognitive science for decades, for example, categorization, reasoning, problem solving, and analogy formation. Others are processes that take place when representations are propagated across representational media, as in the transformation of observations into data and the processing of data to create published inscriptions (e.g., Hutchins, 1995; Latour, 1987). Still others are widely distributed processes that play out in the traffic of

*This work was supported by a grant from the Santa Fe Institute's program on Robustness in Social Processes. We would like to thank the following people for their contribution to this project: Seana Coulson, Hollie Crower, Yrjö Engeström, Gilles Fauconnier, Charles Goodwin, Ed Hubbard, Martin Sereno, Ying Wu, Hsin-Hao Yu, and the participants in the ethnographic study.

**Department of Cognitive Science, University of California at San Diego, e-mail: malac@cogsci.ucsd.edu.

inscriptions and the spread of ideas across scientific communities (e.g., Best & Pocklington, 1999; Fleck, 1935; Galison, 1997; Latour, 1987). In this chapter we will address a different class of cognitive processes that do not easily fit in the categories described above. These cognitive processes occur in the interactions of scientists with one another and with material representations. They are not internal processes that accompany observed behavior; rather we will show that it is useful to see interactions as cognitive processes.

The importance of external representations in reasoning and knowledge construction has been noted by many researchers seeking to understand the nature of the science (Hull, 1988; Fleck, 1935; Latour, 1987; Lynch & Wollgar, 1990). Still, little is known about what actually happens in those moments in which scientists engage one another and their inscriptions. The interactions among scientists and their inscriptions are not only where cognition takes place; these interactions are important cognitive processes in their own right. By focusing on the observable and shared practices of scientists, we provide a bridge between studies that focus on the properties of individual scientists and those that focus on the properties of scientific communities. The cognitive processes that play out in the interactions of human actors with the social and material world are a previously under appreciated domain of cognitive phenomena awaiting exploration by cognitive anthropologists. Because these processes are embedded in the culturally constructed environment of scientific practice, they are at once fundamentally cognitive and especially amenable to anthropological approaches. The careful examination of these interactions reveals action as cognition. We see this approach as a new kind of cognitive anthropology.

Dominance of the Visual

In *The Birth of the Clinic,* Michel Foucault identifies the work of Bichat and his contemporaries as marking the shift from the eighteenth century study of brain pathology, characterized by its "language of fantasy," to the discourse of modern medicine. In the history of science, these eighteenth century anatomists are remembered for their technique of opening human skulls in order to directly observe the brain. With this type of practice, the standard of medical rationality became what could actually be seen with the eye, rather than what had been imagined. The knowledge of the brain was from that point onward linked to perceptual practices:

> [m]edical rationality plunges into the marvelous density of perception, offering the grain of things as the first face of truth, with their colours, their spots, their hardness, their adherence. The breath of the experiment seems to be identified with the domain of the careful gaze, and of an empirical vigilance receptive only to the evidence of visible contents. The eye becomes the depository and source of clarity; it has the power to bring a truth to light that it receives only to the extent that it has brought it to light; as it opens, the eye first opens the truth: a flexion that marks the transition from the world of classical clarity – from the 'enlightenment' – to the nineteenth century (Foucault, 1963, 1973: xiii).

It can be claimed that today we still inhabit this era of *constant visibility*. However, with technological advances, "the artisanal skill of the brain-breaker" (Foucault, ibid.) has been largely replaced with non-invasive techniques of observing the human brain. Currently, one of the most powerful non-invasive techniques is *Magnetic Resonance Imaging* (MRI).

The goal of MRI is to provide detailed static images of anatomical structure of internal body parts, e.g., the brain. This technique uses radiofrequency, magnetic fields and computers to create images based on the varying local environments of water molecules in the body. During a brain MRI scanning session, hydrogen protons in the brain tissues are induced to emit a signal that is detected by the computer, where the signals, represented as numerical data, are converted into pictures of the brain.

In addition to the production of static images of interior regions of the body, such as the brain, the MRI can be applied "dynamically." A new dimension in the acquisition of physiological and biochemical information with MRI is mapping human brain function, or *functional MRI*. Functional MRI (fMRI) is used to produce images of brain functions by visualizing the local changes in magnetic field properties occurring in the brain as result of changes in blood oxygenation. Thus, an fMRI image can show the degree of activity of various areas of the brain: if the image is obtained while a subject is engaged in a particular cognitive task, the image can indicate which parts of the brain are most active in that task. Moreover, employment of different colors on the images can be used to show how the distribution of activity in the brain changes through time. Thanks to this technique, scientists have access to static images or *maps of active processes within the brain.*

Nevertheless, brain images do not function as self-explanatory representations that simply support scientific reasoning and practices. Despite the fact that MRI measurements are made visible, a great deal of interpretative work is required to render the visible images meaningful. This cognitive process of interpretation is the focus of this paper.

To observe this cognitive process, much of which is instantiated in the environment of practice, we record digital video and describe the way in which scientists use fMRI technology to map the human brain. The analysis focuses on the acquisition of specialized fMRI brain mapping knowledge and its organization.

We document how an expert and a novice use a variety of semiotic resources to collaboratively create observable cognitive processes that enable them to identify meaningful entities on the complex visual representation called a "phase map," i.e., the brain image that represents the neuronal activity in the visual cortex (Figure 1). We analyze how representations of *messy* experimental data become organized, meaningful phenomena through the extensive use of gesture, language and material structure distributed across spaces of practice. We also note that scientific knowledge, namely, the capacity to *see* certain concrete and spatially represented natural phenomena, is gradually achieved through the use of dynamic processes of imagination.

Method

The method of cognitive ethnography combines traditional long-term participant observation with the micro-analysis of specific occurrences of events and practices. The cognitive aspects of the observed practice are revealed in the detailed micro-analysis. The two sides of the research, the micro analysis and the larger ethnography, are interdependent.

Our ethnographic study consisted of observations of scientific practices conducted in three laboratories. Fifteen participants were involved in the project. The study took place over a period of nine months, and employed a variety of data collection methods, including direct observation, video recording, semi-structured interviews, and analysis of documents, such as scientific papers, laboratory manuals, and scientific correspondence.

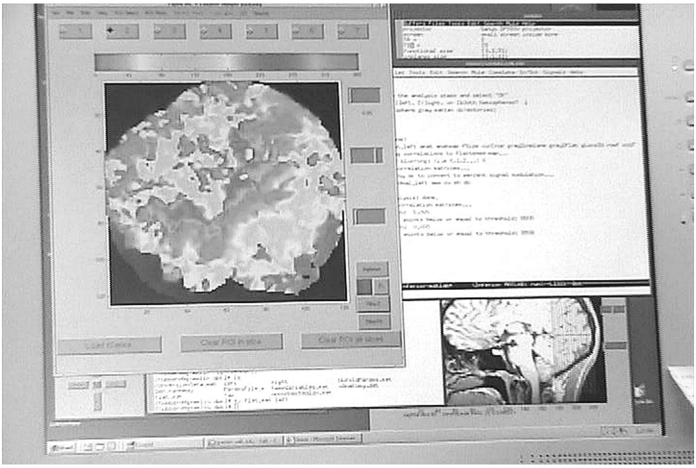

Figure 1. Phase map as viewed on the computer screen.

In order to find out how scientists construct the meaning of complex fMRI images, we perform micro-analysis on digital video recordings of scientific practices. We use these recordings to identify and track the representations that the scientists bring into coordination with the images in the process of making the images meaningful. We use the findings from our long-term ethnography to provide warrants for our analytic judgments concerning how things and events are made relevant to one another in a small instance of scientific practice.

The activity analyzed here, which we will henceforth call "learning to see retinotopy," was selected from eight hours of video recorded during our ethnographic research. This activity was chosen because, as an example of instruction, it makes many of the practices of the laboratory explicit. The passages of the "learning to see retinotopy" presented below are examples of a general pattern observed across actors and laboratories.

Studying Visual Brain Areas

The laboratory, where "learning to see retinotopy" takes place, consists of the principal investigator, two post-doctoral researchers and three graduate students (one of whom is on leave for a year). In addition to the already well-established members of the laboratory, there is a first year graduate student spending a semester in the laboratory in order to acquire knowledge of the methods used in the laboratory. The research of all laboratory members is centered on fMRI measurements of the human brain areas involved in visual processing.

The human brain contains multiple visual areas located in each hemisphere. The general topography and the location of visual areas relative to one another are consistent across individuals. The idea behind the concept of *retinotopy* is that there is an orderly mapping between locations in each retinotopically organized brain area and the locations in visual field. In other words, the retinotopically organized visual areas are considered to be point-to-point copies of the topography of the retina. One should not assume that the representations on the visual cortex look like mental images (or the objects they are representing). Instead, what is of interest are topographic correspondences. That is, the optical properties of the eye project stimuli that are adjacent in visual field to adjacent locations on retina. The connectivity of neurons in visual system maps stimuli from adjacent points on retina to adjacent neurons on early cortical visual areas.

Nearly all visual information reaches the cortex via the primary visual area (V1). V1 is located in the posterior occipital lobe within each hemisphere and provides a precise retinotopic mapping of the visual fields. In the left hemisphere V1, the right half of the visual field is presented, covering 180° of the visual field circle. V1 projects in a topographically well-ordered fashion to V2 (second visual area). V1 and V2 project to numerous visual areas: V3, V3A, VP, V4, V5 or MT, V7, V8, etc. Once the representation of visual stimuli in V1 is established, the other retinotopically organized visual areas can be determined with respect to V1. Successive areas are mirror images of each other.

Since there is significant variability in the size, location and shape of the areas in the visual cortex across individuals, the boundaries of the areas must be determined separately for every individual studied. For this reason, the first year graduate student, the novice (N), has to learn how to chart the retinotopic maps on the human cortex (i.e., to inscribe position and form of visual areas on a phase map, as shown in Figure 2). The practices used in the laboratory to identify visual brain areas are explained by the advanced graduate student, the expert (E). We will analyze a sequence of action through which the novice acquires the capacity to see retinotopically organized areas on the phase map.

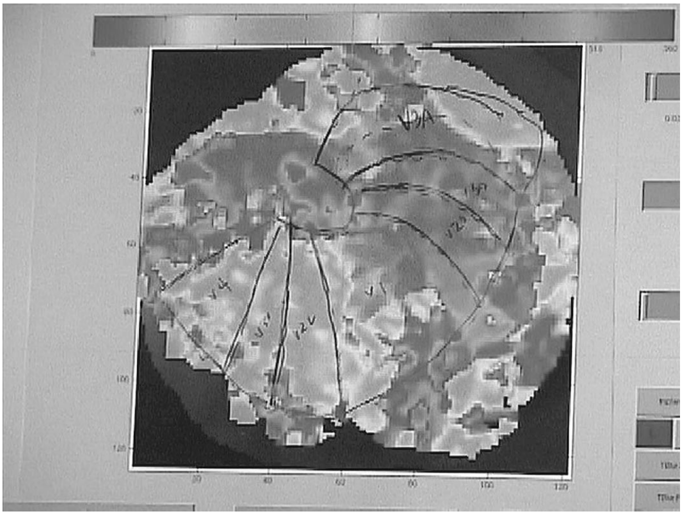

Figure 2. Phase map captured at a later stage of the practice after the expert and the novice had already inscribed on it borders of early visual areas.

The analysis is organized around five excerpts taken from the interaction between the expert, and the novice. The description of the process spans from the expert's stating that identifying visual areas on the brain image is a very complex task that requires knowledge, to the novice's saying that he can see/understand the organization of the visual areas on the brain image. During the entire activity, the expert and the novice are seated in front of the computer screen with a brain image on it (Figure 3). Their interaction first centers around a brain image recorded in response to a stimulus condition called the "expanding ring." This progressively leads into more complex task of defining borders on the brain image configuration recorded in response to "rotating wedge" stimuli. Initially, the expert introduces the procedure by pointing out the difficulty of the task: in order to see the structure on the brain image, one needs specialized training. Next, she tries to explain the relationship between what the experimental subject sees while her/his brain is being imaged, and the neuronal response in her/his visual areas by drawing a chart on a piece of paper. After the chart is constructed, the expert and the novice try to relate the regions drawn on the chart to the structures that should be seen on the brain image. Because of the difficulties that the novice experiences while trying to

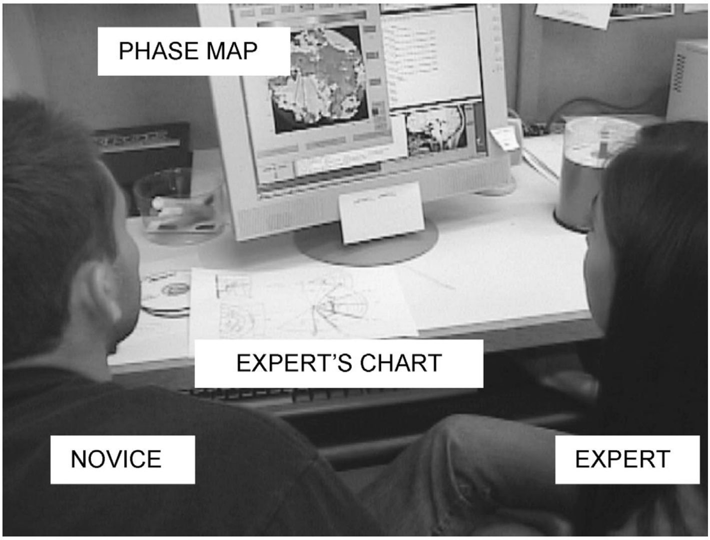

Figure 3. The setting.

make sense of the brain image, the expert attempts to make the structure of the brain map more visible. Finally, the novice says that he can see the organization of the brain image.

Over time, the laboratory has accumulated knowledge about its practice in a laboratory manual. Where the manual describes the process of defining the visual areas on a brain image, it mentions the expert's chart: "Look at M's (the expert here) chart to define visual areas. This is best for expanding rings and rotating wedges. Define the visual areas using the expanding rings and then delineate V1, V2, V3, etc. using the horizontal meridians and data from the rotating wedges. You may want to stop short of V7 or V8 because they are not well defined." In what follows we will show how this abstract description of the action is turned into an effective practice. By looking at the activities of "learning to see retinotopy" in detail, we will describe several instances of action as cognition.

Seeing the expanding ring

During the stages of "learning to see retinotopy," the actors view on the computer screen a phase map (Figure 1). They try to define early visual areas (i.e., V1, V2, V3) on that representation. The phase map is a static

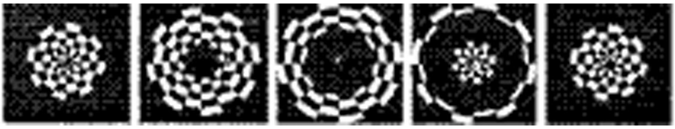

Figure 4. Expanding ring stimulus shown at five moments in time (from Engel et al., 1997).

brain representation that shows the temporal relationship of the data to certain stimuli. On this map, color represents the time at which neuronal activity occurred. The expert and the novice analyze the data from an experiment in which dynamic visual stimuli were used to provoke neuronal responses. The dynamic property of the stimuli produces *waves of activation* such that variation in the temporal phase of the activation is represented by changes of hue in a color map.

The first type of stimuli used in the experiment is called the "expanding ring" (Figure 4, adapted from Engel et al., 1997) – a flickering checkerboard ring that repeatedly expands (or contracts) from the center to the periphery (or vice versa) of the visual field. The subject in the scanner, while presented with the stimuli, is required to fixate the center of the ring. The role of the stimuli is to provoke corresponding neuronal activity in the visual areas. When the stimulus is an expanding ring, within each visual area neurons whose receptive fields are at the center of the visual field (i.e., fovea) will respond earlier than the neurons whose receptive fields are at the periphery of the visual field (Engel et al., 1997). The expanding ring stimulus is important for determining the representation of the fovea in the early visual areas, but not the borders between these areas.

Look to see

The activity captured in the first excerpt from the interaction functions as a preamble to the process of learning to identify the visual brain areas on the brain image. It indicates the beginning of a journey through which the novice acquires the cultural knowledge of the scientific community. This knowledge will allow him to see what exists in the empirical data. In order to find out how such knowledge is achieved, we will closely examine multimodal meaning complexes that develop in the social setting of apprenticeship.

In such complexes the correct unit of analysis is not one semiotic modality, such as speech or gesture taken in isolation, but the entire complex. The meaning of a complex emerges from the interactions among the modalities that include the body, as well as material objects present in the environment. The effects of these interactions are generally not simply additive. Such a meaning complex may be built up incrementally or produced more or less whole, depending on the nature of the components and the relations among them (see Hutchins and Palen, 1998; Alač & Hutchins, in preparation).

When analysing the first excerpt, as well as throughout the text, we'll pay attention to gesture. Even though the study of gesture is a relatively young field of research (e.g., Kendon, 1980; McNeill, 1995, 2000), so far our discussion of gesture has seemed to privilege gestures as phenomena that have meanings in the context of other semiotic modalities. We can try, for example, to identify the meaning of a gesture in respect to the linguistic expression that the gesture accompanies. But this account is inadequate on two counts. First, all of the elements of a meaning complex should be assumed to be of equal status. Second, this usage takes context to be "that stuff I'm trying to ignore at the moment." Rather than trying to identify the meaning of a gesture in a specific context so that we can say "this gesture has this meaning when it is in this context," which gives meaning to the gesture independent of context, we might more accurately say "this gesture-in-this-context has this meaning," which emphasizes the idea that the meaning is borne by the entire complex rather than being borne preferentially by one or another of the constituents of the complex.

This is not to say that we cannot analyze such multimodal complexes. We can. But our expectations for the analysis should not be to isolate the context-independent meanings of the components, or even to isolate the meaning of the separate components in this particular context. Rather it is to identify, if possible, the contribution of the various elements to the accomplishment of the meaning that is borne by the complex of which they are parts.

This stand toward meaning making allows us to treat elements in the social and material environment of action, such as brain images, to be equally important elements of the process. In what follows we will observe how the brain images are brought into the construction of meaning

EXCERPT 1

1 E: So if you look here what can you actually see *((moves closer to the screen))*
2 and it takes quite a bit of training to start and actually see *((the novice moves closer to the screen))* the maps in
3 this noisy data[1]

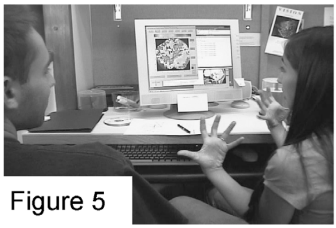

Figure 5

((moving her hands back and forth in front of her, with palms toward her and fingers spread))

complexes, rather than simply functioning as static elements on which various layers of meaning are applied.

By moving closer to the computer screen (line 1), the expert places herself physically closer to the brain image, while directing the novice's attention toward it (the novice's motion in line 2 shows the effect on his attention). The expert's body movement functions as an initial, indexing element in the upcoming construction of the visibility of the image.

At the same time, the expert linguistically invites the novice to look at the data: "if you look here" (line 1). The invitation to look is followed by the utterance "what can you actually see." The utterance is generic: it speaks about the possibility of seeing meaningful structures in the image, rather than the novice's seeing it. The expert is aware that the novice cannot yet see the correspondences between the stimuli and the neuronal response that are to be identified in the brain image. Seeing is not just looking, but perceiving and understanding the image structure. In line 2

[1]Transcription conventions adopted with some changes from Sacks, Schegloff & Jefferson (1979), and Goodwin (1994). // The double oblique indicates the point at which a current speaker's talk is overlapped by the talk of another. = The equals sign indicates no interval between the end of a prior and start of a next piece of talk. (x.x) Numbers in parentheses indicate elapsed time in tenths of seconds. : The colon indicates that the prior syllable is prolonged. ___ Underscoring indicates stressing. ( ) Parentheses indicate that transcriber is not sure about the words contained therein. (( )) Double parentheses contain transcriber's comments and extralinguistic information, e.g., about gesture, bodily movements, and actions. Punctuation markers are not used as grammatical symbols, but for intonation; . Dot is used for falling intonation; ? Question mark is used for rising intonation; , Comma is used for rising and falling intonation.

the expert explicitly expresses the idea that the seeing is not immediate. The collective act of looking/seeing is introduced by saying that the seeing and the identification of "maps in the noisy data" (line 2-3) require training (line 2). The expert introduces the possibility of looking/seeing through the employment of semiotic resources; in this case her body movement and linguistic expression. Simultaneously, she explicitly acknowledges that the seeing depends on the structures external to the specific act of looking, i.e., training.

In line 3 the expert incorporates gesture in the meaning complex to emphasize the active nature of seeing (Figure 5). She moves her hands back and forth between her body and the brain image on the computer screen, with palms toward her and fingers spread. The hands create a trajectory that moves approximately nine times from the map toward her eyes. The gesture participates in the organization of the speech: it highlights elements of the expert speech stream. At the same time, the gesture is also actively involved in meaning construction: it fits with the idea of seeing the maps on the brain image. The idea emerges from the interaction of the gesture with the brain image and the action of seeing expressed linguistically.

Yet, because of its mobility and its location in the space between the brain image on the computer screen and the actors looking at the brain image, the gesture participates in selection of a site of action. It defines the public space between the actors and the brain image as a stage where the transformation of looking into seeing will take place. Rather than focusing on internal cognitive processes that make this transformation possible, we will examine cognitive processes of linkage that will take place in this public space of interaction.

Squeezing the visual field

The meaning complex discussed in the previous section (and shown in Figure 5) suggests the idea that maps are somehow in the noisy data, and they can be seen by those who know how to see them. The seeing is acquired through the relation between the empirical data (represented visually) and training. The training is obtained through participation in the practice as a member of the scientific community (i.e., here the novice learns by actively participating in the analysis of the experimental data).

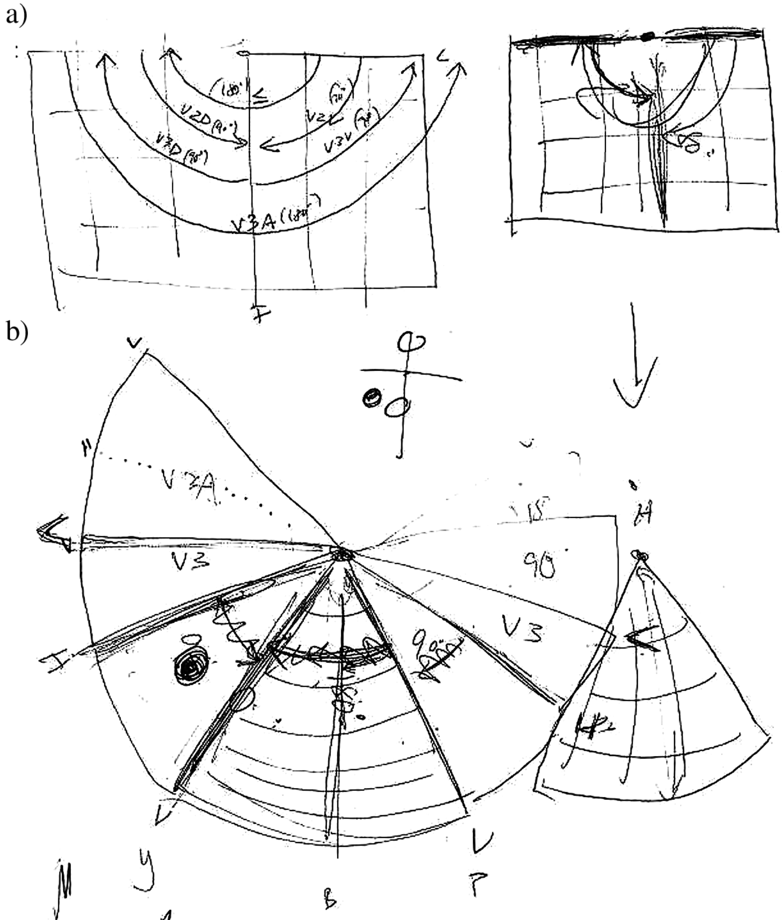

Figure 6 The expert's chart. (a) Visual field representation. (b) Retinotopy space.

At the same time, the knowledge is also deposited in artefacts, such as the expert's chart (Figure 6).

While teaching the novice how to see visual areas on the cortical surface, the expert draws a chart that exemplifies the mapping between the visual field and its projection onto the cortex (Figure 6). The chart can be divided into two parts: the diagram of the visual field (Figure 6a), and the "retinotopy space" (Figure 6b). The retinotopy space represents the cortical organization of visual areas. By the end of the training session,

EXCERPT 2

4 E: So take these two meridians and as just

5 as if you were squeezing them together into the pie shape

Figure 7a

((stretches arms wide apart and then brings them together on the paper))

Figure 7b

Figure 7c

((slowly tilting wrists back and then bringing palms together))

Figure 7

we were shown an already made chart that the expert kept in her office. However, the teaching activity began with a blank piece of paper. This allowed for the chart building to evolve as an interactive and gradual process.

Excerpt comes from the interaction that took place during the chart construction. The expert has drawn the visual field (diagram on the right side of Figure 6a), but not yet its corresponding retinotopy space (diagram on the right side of Figure 6b).

The conceptual relationship between the visual field and its brain representation is being formed in terms of the correspondences between the two parts of the chart – the visual field space and the retinotopy space (as expressed in lines 6 and 7). But how is this achieved?

In order to establish these correspondences, the expert invokes the action of squeezing (line 5). The abstract correspondences between the two structures may be understood in terms of a concrete, physical action. As proponents of cognitive semantics have pointed out, understanding abstract entities and processes is often based upon biological capacities and experiences of functioning in a physical and social environment (Lakoff, 1987; Fauconnier & Turner, 2003). The activity in lines 4-7 is an example of how the participants scale what is invisible and untouchable (i.e., the correspondences between the visual field and the cortical organization of visual areas) to what they can directly experience (i.e., squeezing).

The action of squeezing the visual field representation has no real-world referent. There is no real action wherein the visual space is effectively squeezed and transformed into the retinotopically organized visual areas. In fact, the expert asks the novice to imagine the action: "as if you were squeezing. . ." (line 5). The action must be imagined.

Yet, while the linguistic expression invokes the imaginary action of mapping, the gesture enacts it. Through the expert's use of gesture, that participates in the multimodal meaning complex, the mapping from one representation to the other is performed as a visible and dynamic action. The imaginary process of squeezing, evoked to accomplish the conceptual mapping between the two domains, is instantiated as a concrete, embodied process that unfolds in the environment of practice.

In addition to the gesture, the concreteness of the process depends also on the pencil drawn diagram of the visual field. The drawing is crucial in making the abstract entity, such as the visual field, a concrete one. By being drawn, the visual field acquires the form of a gridded parallelogram. This representation participates in the action of squeezing. The hands are accurately placed on such representation in order to "squeeze" it (Figure 7).

When placed on it, the gesturing hands *glue* onto themselves the visual field representation (the linguistic choice of the verb "to take" used in lines 4 and 6: "so take there two meridians," and "so you took this space," corroborates this interpretation). Thus, the hands stand directly for the lines, i.e., the external borders of the visual field space (notice the rigid form of the hands), which are being squeezed. Through their coordination with the chart the hands carry out the process of change where one representation is transformed into the other. The process reveals two unexpected features of the action. On one hand we see how, through multimodal action, the visual field representation acquires dynamicity. What is perceived as a static representation on the chart is constituted as a dynamic entity. On the other hand we notice the gesture's capacity to carry out different functions. By its proximity with the chart, and the coordination with language, the gesture performs the action of squeezing, and functions as a representation of the visual field.

This excerpt shows how the use of many semiotic resources constructs the mapping between the visual field space and the retinotopy space. In other words, the participants use their hands and eyes to reason (Latour, 1986) about the abstract conceptual transformation. This is an observable cognitive process that is subject to description and an ethnographically informed interpretation. By examining the employment of different semiotic modalities we see how representations are correlated and how the imaginary and concrete action is enacted in the socially constructed environment.

The chart as a collective architecture for seeing

The previous section illustrated how the link between the visual field and retinotopy space is composed from the interaction of different representations, and is realized across a couple of seconds of the multimodal action. However, the construction of correspondences and the manufacturing of the chart contains additional layers of temporality inscribed into it. One such layer is the history of the chart (where the chart is not seen as the specific token constructed on the fly, but a general type).

To find out more about the history of the chart, and to comprehend its role in the expert's understanding of concepts involved in retinotopic mapping, we engaged in a brief written exchange with the expert. The expert's response to a question about the history of the chart was:

> The chart was my idea. I was taught how to interpret the retinotopic maps by someone else in the laboratory (who is no longer here). He basically just showed me the fMRI data on the computer screen (as we saw) and explained it all verbally. I thought it would be easier to understand visually so I put together the chart just to help myself understand. The first version was really messy with many erasures as I was just figuring it all out. So I recopied it in a more orderly fashion (the version that you have). Since then, I have used it several times to teach others. People have found it useful and have asked me for copies.

Her response reveals the temporality of the knowledge inscribed in the chart. The chart is a distillation of prior experience. The expert constructed the chart after she was taught how to interpret the retinotopic maps. Yet, the fact that the expert's chart is being referred to in the laboratory manual, indicates that the chart, by being a ratified element of the laboratory practice, contains not only M's voice, but voices of the entire laboratory community.

The expert's response also shows that the chart was not simply constructed to transmit her knowledge to others. On the contrary, the chart was made in order to help her understand the task: "I put together the chart just to help myself understand." In this sense, the chart is a representation constructed in order to comprehend the other representation (i.e., the visual representation of the experimental data). The chart functions as what Goodwin calls an *architecture for seeing* (personal communication) or *structure of intentionality* (1994: 609). The fact that the construction of the chart was not just a simple translation of some internal structure (i.e., the knowledge previously acquired by the expert) into an external one, but a loop-like process, where the construction is gradually modified with respect to the feedback given from the chart's structure, is crucial. This can be seen when the expert says: "The first version was really messy with many erasures as I was just figuring it all out. So I recopied it in a more orderly fashion [. . .]." Accordingly, the chart inscribes into its structure the history of the collective laboratory endeavor, as well as the temporality of the expert's actual manufacturing of the artifact.

Notice also that while the information was given to the expert verbally, she re-represented it visually: "He basically just showed me the fMRI data on the computer screen [. . .] and explained it all verbally. I thought it would be easier to understand visually." This adding of an intermediate structure between showing and seeing, and by re-representation of information in a particular format made the task cognitively easier to manage (Kirsh, 1995; Clark & Thornton, 1997). The usefulness of the chart in constructing of understanding will be illustrated through the analysis of the excerpt which follows.

Correlating chart, hand, and image to make the unknown known

Equipped with the chart, the actors can begin to map the territory represented on the brain image (see Figure 8).

As the participants' discursive action described in excerpt 1 shows, the understanding of the brain image structure depends on the competency to see such structure. The idea that visual representations are inseparable from such competencies is not new (e.g., Wolgar & Lynch, 1990). Nonethe-

EXCERPT 3

8 So probably this is the center *((touches the screen with her index finger))*

9 right here *((takes the sheet of paper with the drawings and moves it clockwise and points to what represents the fovea int the "retinotopy space"))*

10 And when we look at this map it looks something like that *((picks up the paper and holds it next to the computer screen))*

11 So V1 is gonna be in the center *((briefly traces the borders of the V1 representation of the retinotopy space on the paper with her index and middle finger))*

12 it's gonna be this pie shape it's probably covering approximately this area

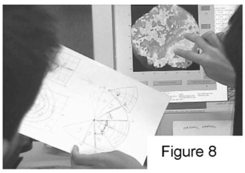

Figure 8

((carefully places her index and middle finger on the "center" of the phase map on the computer screen and traces the imaginary borders of the V1 representation. Repeats the movement six times))

13 Ok?

Figure 8

less, the issue of how the brain representation is linked to competence in seeing its organization is as yet unresolved. The following analysis shows how by looking at multimodal elements of discursive action we learn about the interaction between the unknown and the known.

The interaction and linking between the unknown and the known is visible in the activity in lines 8-13. The goal for the participants is to determine where the brain area V1 (which is labelled on the chart) should be located on the brain image. The analysis of the passage allows for an exploration of how different structures present in the environment -some permanently, others only transitorily – are assembled and linked together to construct the novice's seeing of the brain image.

The expert's language and gesture show the expert's own uncertainty about the organization of the brain image in respect to the chart. When the expert in line 8 indicates the center of the phase map, she uses the word "probably" ("so probably this is the center"), while her indexing of the center on the expert's chart in line 9 is accompanied by the expression that indicates certainty ("right here"). Similarly, in lines 11-12, when the expert indicates the position of V1 by tracing its border on the brain image, she uses the words "probably" and "approximately" (line 12 "It's gonna be this pie shape it's probably covering approximately this area"), and chooses to speak in the future tense ("it's gonna be" in lines 11 and 12).[2] Yet, while in excerpt 1 the mapping process was suggested through speech, here the mapping is largely achieved through the use of non-linguistic semiotic forms.

In order to see how the meaning of the retinotopic map and the identification of its center is achieved, we'll pay careful attention to two elements of the process: 1) knowledge organization of the known (that has to be inscribed into the unknown) and 2) the mapping processes between the known and the unknown. Curiously enough, both elements are instantiated in the external, material and social, world of practice.

As already noted in the previous section, the knowledge organization and its form are laid out in the chart (Figure 6).[3] In order to transform the brain image into what is for both participants a known domain, the structure of the knowledge existing in the chart has to be mapped onto the brain image. In this process the chart functions as a representation of a schema through which the brain image is understood.

In line 8, the expert points to the probable location of what stands for the fovea on the brain image. Then, in line 9, she indexes its representation on the chart. Note that she, counterintuitively, first points toward the unknown and only then toward the known structure. This pointing sequence indicates that the pointing toward the brain image is a general process of attention directing, rather than an act of indicating a particular form on the image. But importantly, and this will be described in more detail in what follows, this pointing also functions as a preparatory stage for an activity where the known structure will be *superimposed* on the unknown one. In this sense the pointing in line 8 functions as a *taking out* of the structure existing in the image in order to further elaborate on it in the chart. The gesture translates the image into another, better known structure – the chart. To make the mapping between the two representations more effective, the expert rotates the chart (line 9) so as to prepare for a more straightforward tightening of the relationship between the two.

---

[2]The uncertainty that the words in lines 15 and 16 express depends as well on the fact that the actors are presently looking at the "expanding ring" map where the borders between different areas are not clearly defined by the colors on the map.

[3]This is, obviously, achieved once the actors know how to interpret the structure of the chart.

The role of the chart and the mapping of various representational forms are further developed in lines 10-12. In line 10 the expert picks up the paper and holds it next to the computer screen. The physical act of placing the chart closer to the screen and holding it in such a position invites the interpreter to see the structures present in the two representations as comparable. Next, a particular element of the chart, i.e., the representation of the visual area V1, is mapped from the chart onto the brain image. A crucial part of the action is played out by gesture. In line 11 the expert places her right hand onto the chart and briefly traces with her index and middle finger the borders of the V1. In line 12 the hand is lifted from the chart, moved toward the brain image, and carefully placed onto the image where a gesture of the similar form is executed.

The action is not simply an execution of two indexical gestures (one over the chart, and the other over the phase map). Rather, the expert performs a complex, multimodal action that progressively unfolds through time. Since the borders of V1 are already drawn on the chart, the gesturing in line 11 does not merely *draw* the structure on it. Instead, the gesturing hand, via the touching of the chart, *selects* the already existing form and position (with respect to other visual areas) of V1, quickly enacts it on the chart, and carries it over onto the phase map. In other words, the gesturing over the chart is performed not so much to establish the new structure (as in the case of the gesturing over the phase map), but to *pick up* the existing one and *transfer* it over to the phase map. On the other hand, the gesturing on the image (line 12), rather than enacting what already exists in the representations, superimposes onto the image a new, up to this point invisible, form transferred from the chart. The lines (i.e., the borders of the V1) that have to be deduced by observing the change in color (representing a different time of neuronal activation) are now *inscribed* on the brain image through the gesturing act.

Notice that just by looking at the image the novice cannot yet infer the position of the V1 borders. However, he can observe the solution to the problem, as well as the process of its production, before his eyes. He sees the lines, and hence the V1 borders, by observing the expert's gestures on the image. He also participates in the way in which such lines are generated. He views the coordination of the chart and the brain image while such coordination is enacted in the shared space of action. The interaction of the expert's gesture with the visual representations inscribes the (imaginary) lines onto the image and produces the mapping between the representations.

For the expert, on the other hand, acts of mapping are not complete until she has performed the action. She knows she is going to inscribe an area on the brain image of approximately pie wedge shape, of approximately this location, and this size. She knows this because she has an internalized knowledge of what is represented on her chart. But until she does the inscribing action on the brain image she probably does not *see* the border of V1 on the image. Just as thinking out loud can be both communication and thinking, acting can simultaneously serve communicative and cognitive functions. Cognition and communication are not mutually exclusive categories.

At this point one could ask: why is it important to recognize that the gesturing in lines 10-12 is not only indexical in character? And why do we need to speak about the complex gesture of interaction rather than single gestures taken in isolation?

It is essential to notice that the global meaning of the activity (i.e., the mapping between the chart and the image, and the carrying over of the V1 form from one representation to the other) emerges from concatenation of different elements of the process. If we were to describe what is going on in lines 10-12 in terms of the two representational gestures (over the chart and over the brain image), our description would not account for the overall meaning of the activity. In order to account for it we need to consider such gesturing as component of the complex process of interaction. For example, in lines 11-12 the meaning of the gestural action is closely linked to the type of representational form on which such action is performed (the chart vs. the brain image), as well as the participants' knowledge of such form (the chart as a known structure vs. the brain image as an unknown structure).

The same is true for the expert's chart. While it functions as a schema for the understanding of the brain image, it doesn't produce meaning on its own. Instead, like the expert's gesturing, the chart participates in the multifaceted action. Rather than isolating the chart from the rest of the action, we are interested in the process of mapping the chart to other representational structures. Through such mapping, elements of the chart become *dynamic*, *transportable* structures. In comparison to the static form of the chart (when the chart is considered to be a self standing, isolated form), the compound multimodal structure is less permanent, but a more powerful, dynamic process. Because this process is largely instantiated in the environment, we can see it and examine its unfolding through time in detail.

But importantly, the participants in action can do so as well. The fact that the gesturing in lines 10-12 is not just indexical, points out that the participants are dealing with an external cognitive process. While one could argue that the pointing in lines 8-9 and the bringing closer of the two representations in line 9 surely require an internal process of mapping in order to coordinate the two structures, this is not entirely true for the gesturing in lines 10-12. The gesturing does not only indicate the structures that have to be linked. By coordinating with other semiotic means, gesturing performs the linkage. It has often been pointed out that this process of mapping and integration across different domains is characteristic of human cognition (e.g., Fauconnier & Turner, ibid.; Koestler, 1964; Mithen, 1998). Here we see such a process developed in the external environment shared by the participants. Noticing the complex process of interaction allows us to speak about action as cognition.

Seeing the rotating wedge

After the center of the map (i.e., the representation of the fovea) has been established on the brain image that was induced with the expanding ring stimulus, the actors trace borders between the visual areas. This is accomplished on *rotating wedge phase map*. The structure of the map has to be related to the *rotating wedge* stimulus that caused it: a flickering wedge that rotates slowly around the fixation point (Figure 9). Like the expanding ring, the rotating wedge stimulus ideally provokes the alternation frequency of the neuronal response that varies in the temporal phase.

The sum of the neuronal responses creates what scientists call "the traveling wave of neuronal activity." The traveling wave is a collective property that appears in populations of neurons in retinotopically organized visual areas. Because V2 contains a mirror image of the visual field representation in V1, the direction of the neuronal response or the trav-

Figure 9 Rotating wedge stimulus shown at five moments in time (from Engel et al., 1997).

eling wave in V2 will be reversed with respect to V1, and it will reverse direction again in V3.

The periodic structure of the neuronal response is represented on the phase map by a range of color. For example, the rotating wedge stimulus provokes a neuronal response in V1 such that the activity at locations containing neurons whose receptive fields are further along the direction of rotation is delayed with respect to locations containing neurons whose receptive fields are near the stimulus starting position (Engel et al., 1997). Consequently, the neuronal groups whose response was delayed will be marked with a different color than those whose response occurred prior to them. By tracing the change in color, the scientists are able to identify borders of adjacent visual areas, which appear as reversal in the color sequences.

Point to think and see motion

During the following two excerpts the borders of the visual areas had already been traced on the brain image (Figure 2). Nonetheless, the novice expresses his uncertainty about the location of the borders. To lessen the problem the expert tries to find a better view of the data (she *rotates the color map*[4]) so that she can more convincingly justify the position of the borders on the brain map.[5]

[4]She changes the mapping between colors and the time of neuronal response so that, for example, a response phase that was represented in the original color map by red may become yellow in the rotated color map.

[5]This suggests that the visual properties of the experimental data representation are of importance for the categorization and understanding of the empirical data. If that was not the case, the manipulation of the color could not help in identifying the borders between the areas. After the manipulation, the data remain the same, but the alteration of the way in which they are presented can influence their understanding (i.e., certain features in the data can or cannot be noticed).

EXCERPT 4

14	E:	So now it's going from re::d (.) red to pink to blue (3.5) and maybe it goes out to red to pink again and back to blue ((while pointing different colors on the map))
15		(and up there where it breaks down) ((briefly points again onto the upper portion of the map))
16		((Mumbles while going over the color scheme)) ((points back and forth between different colors on the map))
17		((she raises her voice, and her speech becomes clearer when she pronounces the name of the second and subsequent visual areas)) V2 V3 (1.0) V4 ((puts her hand down))
18		So ok so so one theory is that hahaha ((silently laughs)) ok V1 (.) so pink to blue is V2 (2.0) ventral then blue to pink (1.0) is V3 ventral and then pink back to blue is V4(([points different colors on the map])
19		((puts her hand down, and turns toward the novice))
20		That's my best guess based on the data
21	N:	((tries unsuccessfully to take the floor))
22	E:	Even though it's very unclear
23	N:	So V1, V2v, V3, and V4 ((while pointing with the pencil on the map))

The complex cognitive task of identifying the change and reversal of color sequence to justify the position of the borders can be simplified through an engagement of a variety of external material representations. The representations are permanent as well as transitory in character. One such representation is a horizontal array of colors situated above the brain image on the computer screen (Figure 5). The array indicates the phase of the neuronal activity for voxels being active as a response to the stimuli. The expert can directly orient to the colors displayed on the color array. In order to identify change and reversal of colors on the image, the expert can glance at the horizontal color array. She can then match the pattern of colors displayed on the array with the colors present on the brain image.

This pattern matching activity is embedded in a larger meaning complex that includes elements in gestural and linguistic modalities. The expert starts identifying the sequence of colors on the image by placing her index finger on the red color region (line 14). She pronounces the word "red" by prolonging the sound "e." While she pronounces it she can look at the horizontal color array and notice that the red should be followed by pink and blue color. Consequently, she can try to identify the same sequence of colors on the image. Prolonging the sound "e" in "red" keeps the idea of "red" as an active representation in the environment. In such way the memory of the starting point of the color identification is situated in the environment of practice.

After identifying the change from red to pink to blue, the expert pauses for 3.5 seconds. While pausing, she keeps her index finger on the same position on the map. The gesture assumes a function similar to the prolongation of word pronunciation. By keeping her finger in this constant position, the expert can look for a consequent change in color. The index finger is an external memory tool which actively participates in inference making. Through the extensive use of external representational resources, inference is largely produced as a pattern matching process.

In line 16 the expert tries for a second time to list and point out the colors in order to define the visual areas. She again indexes the map, but her voice is very soft. She seems to be mumbling to herself, using speech as self-regulatory process, rather than talking to the novice. While mumbling she points back and forth between different colors on the map. In line 17 her mumbling turns into clear speech, and her pointing becomes linear. She starts enumerating the visual areas from V2 up.

By focusing on the expert's linguistic performance and the use of gesture we see action as cognition. We observe how the expert constructs her understanding of the experimental data configuration through her interaction with the information present on the computer screen. Making sense of the image is progressively built through action, where thinking involves the employment of external elements of thought.

In line 18 the expert, for the third time, lists colors on the map, and points out the borders between the visual areas. By premising her pointing with the phrase "one theory is that" followed by a sotto voce laugh, she indicates the difficulty in identifying the areas in a straightforward way. Consequently she turns toward the novice to check if he saw what she saw and what she had indicated to him (line 19). Next, the expert adds that the identification of the areas is still somehow tentative. This can be seen in line 20 where she states, "That's my best guess based on the data," and again in line 22: "Even though it's very unclear."

Likewise, the excerpt that follows will illustrate how, in an attempt to make the border clearer, the expert rotates the color map. While observing the new configuration on the brain image, the expert expresses her appreciation of it. After saying that the map looks nice, she describes the change in color in the following manner: "Actually this rainbow looks nicer now, doesn't it. Hhhh ((quietly laughs)) the V1 (goes from) blue to purple to red." However, the novice disagrees and shows his preference for the configuration of color present at the view that the expert had previously explained to him: "I like the blue to green the other combination." A configuration is *nice* or *likable* when it's easily matched with the horizontal array of color above the image. In order to demonstrate that the current configuration has its advantages, the expert tries to point out the change in colors more closely:

EXCERPT 5

24	E:	Again so you have now it's yellow orange right? ((points with her index finger and traces the lines over the borders of the phase map))
25	N:	Mhm
26	E:	Mmmm ((hesitates)) then it goes out to purple and then back to orange and
27		then out to purple again ((still pointing and tracing the lines)
28	N:	I actually I can see that now Hhhh
29	E:	You kind of see some of that intermediate (stuff where) it goes from
28		orange to red to purple right? ((points on the computer screen)
29	N:	(I guess that is better)
30	E:	Yeah ok. Right. ((the expert takes her hand away from the computer screen, claps over the table and quietly laughs))
31		If you (would) believe me (it would be) very nice haha ((laughts))
32	N:	No I see it I see what you are saying I see what you are saying
33	E:	Haha ((laughts)) Ok.

This passage is another example of how the novice has to *see* the change in color in order to understand why the borders between different visual areas are placed at certain positions. Once again, the act of teaching the novice to see retinotopic maps is performed through the coordination of different semiotic modalities.

In lines 24, 26, 27, 29, and 30 linguistic expressions and gestures are used in coordination with the phase map. The excerpt illustrates that language and gesture are not limited to the description of what already exists in the brain image. They can supplement it with layers of meaning that are dynamical and fictive. Such elements can be crucial for the execution of cognitive task. Whereas in excerpt 2 illustrated gestural involvement in the production of imaginary and dynamic entities, here such aspects are mainly generated through linguistic usage.

In lines 26 and 29, in order to describe the change in color, the verb "to go" is used. Even though the colors on the phase map do not move, i.e., the phase map is factively a static representation, the subjects describe it in terms of movement. Leonard Talmy (e.g., Talmy, 2000) calls this phenomenon *fictive motion*. In fictive motion one deals with a non-veridical event in which a static entity is construed as being in motion. For example, we can linguistically depict the form, orientation, or location of an object in terms of a trajectory over the object's extent.

As Talmy has pointed out, fictive motion can be expressed linguistically as well as perceived visually. In fact, the subjects almost seem to be perceiving motion on the static phase map. For example, in line 29 the expert states: "you kind of see some of that intermediate (stuff where) it goes from…" By moving their gaze over the map, the participants act as if they see the progressive change in the static scenery. By expressing motion, the language inscribes dynamicity onto the map and enhances the embodied experience of the visual representation.

The opportunity to read the experimental data in terms of a dynamic change in colors allows the actors to process the data in a way that is particularly suitable for human cognition. If the same data were represented numerically, it would be very difficult, or perhaps impossible, to determine the propagation of the *traveling wave of neuronal activity*, described here in terms of fictive motion.

While the expert expresses the motion verbally, she actively gestures towards the computer screen. Her gestures are coordinated with her talk. As she mentions different colors, her gesture briefly traces the change in color on the map. However, the gesture is primarily oriented towards outlining the boundaries of the visual areas. Hence, the gesture communicates, but only in passing, what is linguistically expressed, as the expert's index finger moves across the visual areas indicating color. At the same time, it indicates the boundaries of the visual areas. The expert does not explicitly talk about the boundaries. The novice is expected to infer the existence of such boundaries from the linguistic description of the phase map configuration. And, what he is expected to infer, i.e., the location of the boundaries, is indicated by the expert's gestures.

In line 28 the novice assures the expert that he can see the form and location of different visual areas: "I actually I can see that now." However, the expert insists again in highlighting what the novice should see (lines 29-30). In line 31 the novice acknowledges that this configuration of colors is more informative: "I guess that is better." His position has changed as a result of the expert's use of different linguistic and gestural representations (lines 24-30) employed to produce visibility of the image.

In line 31 the expert, jokingly combines the idea of believing and seeing: "If you (would) believe me (it would be) very nice." Her utterance highlights the idea of seeing as situated in the interaction with others (i.e., the novice should believe her). Since her voice *speaks for* the laboratory as well as the larger neuroscience community, believing what she says situates seeing in the enormously complex structures of cultural knowledge.

The novice in his reply (line 32): "No I see it I see what you are saying I see what you are saying," tries to convince the expert that he does see the position and the form of different visual areas, as well as understands why such areas can be characterized as such. In doing so he first states that he *sees* and then repairs his utterance twice by saying that he *sees what the expert is saying*. His utterance shows again how his seeing is the product of the expert's *saying*. And the expert's *saying* involves a coordination of multiple structures built through the situated interaction of learning to see.

Interestingly, the novice's expression is representative of the idea that underlies the fMRI practice in general. His expression can be read as not simply saying that he understands what is being said, but that he understands through seeing. The fMRI technique, by spatially identifying areas where particular brain processes take place, allows researchers to infer the nature of specific cognitive processes. Thus, it can be said that the fMRI is based on the metaphor of "seeing is believing" (Kutas et al., manuscript). At the same time, it should not be overlooked that *believing* is not simply *looking*, but *seeing*. Seeing is instantiated through observable cognitive processes that coordinate different embodied and cultural representations.

Discussion

We began the paper by asking how brain researchers accomplish the meaningful interpretation of digital brain images. The answer is complex. Our description of learning how to recognize retinotopically organized visual areas indicates that the knowledge necessary for the individuation of meaningful entities within a complex representation, such as a retinotopic phase map, is constructed by recourse not only to the digital image of brain activation, but by coordinating the brain image with other cultural and embodied representational means distributed across the environment of practice (Hutchins, 1995).

We develop the idea of action as cognition. We intend this idea to be an alternative to, and a challenge to, traditional views that locate all cognitive processes inside the heads of individual actors. We showed how brain researchers perform acts of scientific imagination in coordination with observable representations. A diagram of the visual field provides an anchor for an imagined visual field object which is then squeezed with the hands to make its shape match the shape of the retinotopy space. This is an embodied enactment of a complex transformation that accounts for the topological relationship between the visual field space and the retinotopy space. This enactment is fictional in the sense that it represents the transformation as a process, squeezing, when, in fact, the transformation is produced in the visual system by the structure of neural circuitry. The researchers use their hands to highlight a shape in the retinotopy space and then move the hand carrying the shape to the brain image. This is the manual enactment of the imagination of shape similarity across the two material representations. The researchers use their fingers to inscribe the borders of the brain areas on the phase map representation. This is the enactment of the imagination of clear distinctions in a complex and continuous display. These actions constitute thinking for the researchers. The researchers both produce the meaning of the scientific inscriptions and understand the meaning that is produced by performing these actions. They employ their bodies to find out how the two diagrams on the chart are related to each other, how elements from chart and the brain image are associated, and how the brain image is organized.

Gesture is an especially important semiotic resource in interactions that concern spatially organized displays. By analyzing gesture as a feature of action as cognition, we notice that gesture not only conveys or encodes meaning. Gesture often ties together different representations. Gesture can link distinct representations into larger schematic units, highlighting potential but otherwise unmarked dependences among the representations. For example, the "squeezing the visual field" gesture builds a link between the sketched visual field representation and the sketched retinotopy representation by creating an imaginary dynamic process. The inscribing V1 on the brain image gesture picks up the V1 representation from the expert's chart and transports it to the brain image. The gesture thus highlights relations among these representations that can be imagined but are not at all obvious. Importantly, these relationships and linkages are performed in the public space of action. The problem of linking disparate domains into composite conceptualization is one of the central questions in cognitive science (for discussion see Fauconnier & Turner, ibid.). Here we see such cognitive process being executed in the external, shared world.

A complete account of cognition in scientific research (or any other setting) would include careful descriptions of both internal cognitive processes and the processes that occur in the interactions of individuals with one another and with the material environment of action. However, no existing technology or method makes it possible to know in detail what is happening inside the heads of people when they are engaged in the sorts of scientific activities that are the focus of this paper. Hypotheses concerning what people do with their minds must be grounded in a specification of the nature of the cognitive tasks that are faced by individuals in this setting. The specification of the cognitive tasks that are actually engaged by individuals in any setting can only be achieved by carefully examining the organization of real action in that setting.

For example, the expert must be able to use the patterns of color on the phase map image to locate the boundaries of the visual area, V1. She must be able to search the phase map for regions where the color progression (defined by the color map) reverses. She must then somehow construct imagined lines along which the reversal of the color map occurs. These are some of the cognitive tasks demands of the observed activity. The expert meets these task demands by using her finger to trace provisional or tentative lines that she can then *see* and evaluate for their adequacy as boundaries between regions in which the organization of color is reversed. The cognitive task of determining the boundaries is not simply expressed in gesture. Producing the tentative gestures is part of the process that accomplishes the cognitive task. It is therefore properly seen as an observable and external aspect of a cognitive process.

There are two important aspects to notice about this task. First, this is not the sort of cognitive capability that has traditionally held center stage in cognitive science or psychology, yet it is absolutely central to the practices observed here. Second, while one may speculate that visual imagination is involved, the internal knowledge and the cognitive processes that realize these acts of visual imagination are, at present, unknown.

Given the simultaneous presence of two kinds of cognitive processes: those that are hidden inside of actors, and those that are available for observation in the interactions among actors, and given that the former have received intense study while the latter have been for the most part overlooked, we choose to examine action as cognition. This seems to us to be the natural province of cognitive anthropology: enacted cognitive processes that are intrinsically constituted in the coordination of external cultural representations.

Conclusion

By analyzing action as cognition we are constantly reminded that the social and material worlds are intrinsic elements of conceptualization and inference production (Goodwin, 1994; Hutchins, 1995). This is true not only for everyday, "simple" practices, but also for complex scientific activities. Although scientific practices initially appear quite abstract, the objects of knowledge that these activities produce are built through mundane processes of situated human interaction.

Our analysis concerns the practice of reproducing the knowledge of retinotopic mapping. We examine several excerpts from the practice in order to illustrate the trajectory that took the novice from *looking* at the image to *seeing* the structures in it. Since *seeing* requires training (as was explicitly pointed out by the expert), it involves coordination of different processes and representational structures existing not only in the individual minds of the practitioners, but also in immediate environment of their practice.

In order to acquire knowledge and understand the structure of the brain image, the novice learns to see through the *expert's saying*. We saw how *expert's saying* embraces structures and processes that span from gestural actions to institutionally recognized repositories of the cultural knowledge of the scientific community. Since knowing is social, then seeing is as well a collective, cultural, accomplishment, distributed across practitioners and material world.

REFERENCES

ALAČ, M. & E. HUTCHINS

In prep. 'Finding Structure and Fiction in Scientific Practice.'

BEST, M. & R. POCKLINGTON

1999 'Meaning as Use: Transmission Fidelity and Evolution in Netnews.' *Journal of Theoretical Biology* 196: 385-395.

CLARK, A. & C. THORNTON
1997 'Trading Spaces: Computation, Representation, and the Limits of Uninformed Learning.' *Behavioral and Brain Sciences* 20: 57-90.
ENGEL, S.A., G.H. GLOVER & B.A. WANDELL
1997 'Retinotopic Organization in Human Visual Cortex and the Spatial Precision of Functional MRI.' *Cerebral Cortex* 7: 181-192.
FAUCONNIER, G. & M. TURNER
2002 *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books.
FLECK, L.
1935 *Enstehung und Entwicklung einer Wissenschaftichen Tatsache. Einführung in die Lehre vom Denkstil und Denkkollektiv*. Basel: Schwabe. (Engl. trans. *Genesis and Development of a Scientific Fact*. Chicago: The University of Chicago Press, 1977).
FOUCAULT, M.
1963 *Naissance de la Clinique*. Paris: Presses Universitaires de France. (Engl. trans. *The Birth of the Clinic: An Archeology of Medical Perception*. New York: Random House, 1975).
GALISON, P.
1997 *Image and Logic: A Material Culture of Microphysics*. Chicago: The University of Chicago Press.
GOODWIN, C.
1994 'Professional Vision.' *American Anthropologist* 96(3): 606-633.
1995 'Seeing in Depth.' *Social Studies of Science* 25: 237-274.
2001 'Practices of Seeing Visual Analysis: An Ethnomethodological Approach.' In T. Van Leeuwen & C. Jewitt (Eds.), *Handbook of Visual Analysis*. London: Sage.
2002 'Pointing as Situated Practice.' In S. Kita (Ed.), *Pointing: Where Language, Culture and Cognition Meet*. Hillsdale, NJ: Lawrence Erlbaum Associates.
HULL, D.
1988 *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*. Chicago: The University of Chicago Press.
HUTCHINS, E.
1995 *Cognition in the Wild*. Cambridge, MA: MIT Press.
HUTCHINS, E. & L. PALEN
1998 'Constructing Meaning from Space, Gesture and Speech.' In L.B. Resnick, R. Saljo, C. Pontecorvo & B. Burge (Eds.), *Discourse, Tools, and Reasoning: Essays on Situated Cognition*. New York: Springer Verlag (NATO ASI Series F, vol. 160).
LAKOFF, G.
1987 *Women, Fire, and Dangerous Things*. Chicago: The University of Chicago Press.
LATOUR, B.
1986 'Visualization and Cognition: Thinking with Eyes and Hands.' *Knowledge and Society* 6: 1-40.
1987 *Science in Action*. Cambridge: Harvard University Press.
LYNCH, M & S. WOOLGAR (EDS.)
1990 *Representation in Scientific Practice*. Cambridge, MA: MIT Press.

KENDON, A. (ED.)
1981 *Nonverbal communication, Interaction, Gesture: Selections from Semiotica*. The Hague: Mouton de Gruyter.

KIRSH, D.
1995 'The intelligent use of space.' *Artificial Intelligence* 73: 31-68.

KOESTLER, A.
1964 *The Act of Creation*. New York: Macmillan.

KUTAS, M., K. FEDERMEIER, R. SCHUL & J. KING
Manuscript on the inference problems involved in *brain mapping*.

MCNEILL, D.
1995 *Hand and Mind: What Gestures reveal about Thought*. Chicago: The University of Chicago Press.

MCNEILL, D. (ED.)
2000 *Language and Gesture*. Cambridge: Cambridge University Press.

MITHEN, S. (ED.)
1998 *Creativity in Human Evolution and Prehistory*. London: Phoenix.

SACKS, H., E. SCHEGLOFF & G. JEFFERSON
1979 'A Simplest Systematics for the Organization of Turn-taking in Conversation.' *Language* 50(4): 696-735.

TALMY, L.
2000 *Toward a Cognitive Semantics*. Cambridge: MIT Press.

TOOTELL, R.B.H., N.K. HADJIKHANI, J.D. MENDOLA, S. MARRETT & A.D. DALE
1998 'From Retinotopy to Recognition: fMRI in Human Visual Cortex." *Trends in Cognitive Science*, Vol. 2(5), May.

WANDELL, B.A.
1999 'Computational Neuroimaging of Human Visual Cortex.' *Annual Review of Neuroscience* 22: 145-173.