Title
How brain images reveal cognition : an ethnographic study of meaning-making in brain mapping practice

Permalink
https://escholarship.org/uc/item/68v8z4sn

Author
Alač, Morana

Publication Date
2006

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

How brain images reveal cognition: An ethnographic study of meaning-making in brain mapping practice

A dissertation submitted in partial satisfaction of the

requirements for the degree of Doctor of Philosophy

in

Cognitive Science

by

Morana Alač

Committee in charge:

Professor Edwin Hutchins, Chair
Professor Lisa Cartwright
Professor Seana Coulson
Professor Gilles Fauconnier
Professor Charles Goodwin

2006

The dissertation of Morana Alač is

approved, and it is acceptable in quality and form

for publication on microfilm.

Chair

University of California, San Diego

2006

DEDICATION

To Josipa

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

This dissertation consists of six chapters. Four of the chapters have been published. The dissertation author is the primary investigator and author for all four published chapters.

Chapter 2, in full, is a reprint of Morana Alač: "Negotiating Pictures of Numbers", *Journal of Social Epistemology*, 18:2, 199-214, 2004. http://www.tandf.co.uk

Chapter 3, in full, is a reprint of Morana Alač: "From Trash to Treasure: Learning about the Brain Images through Multimodality", *Semiotica*, 156-1/4, 177-202, 2005.

Chapter 4, in full, is a reprint of Morana Alač & Edwin Hutchins: "I See What You are Saying: Action as Cognition in fMRI Brain Mapping Practice", *Journal of Cognition and Culture*, 4:3, 629-661, 2004. Permission to reprint was granted by the co-author.

VITA

1997	B.A. (Hons.) Philosophy, University of Bologna, Italy
2002	Ph.D. in Semiotics, Department of Communication, University of Bologna, Italy
2004	M.S. Cognitive Science, University of California, San Diego, CA.
2006	Ph.D. Cognitive Science, University of California, San Diego, CA

ABSTRACT OF THE DISSERTATION

How brain images reveal cognition:

An ethnographic study of meaning-making in brain mapping practice

by

Morana Alač

Doctor of Philosophy in Cognitive Science

University of California, San Diego, 2006

Professor Edwin Hutchins, Chair

The study investigates the nature of multimodal communication as a process situated in socially organized and culturally rich environments of practice. To do so it describes ways in which neuroscientists make meaning of the visual fMRI evidence. It examines everyday laboratory practices (e.g., how a novice learns through interaction with an expert and the material world of practice, and how neuroimaging data are analyzed and presented to the larger scientific community), to describe how practitioners engage with digital brain images while they coordinate a variety of multimodal semiotic means such as gesture, talk, gaze, and body orientation. The study involves traditional ethnographic observations and digital video and audio recordings of practices conducted in three laboratories in the Cognitive Science department at UCSD and the Salk Institute. By acknowledging the mutual dependence of communication, signification, and inference the study also advances claims about two theoretical currents in the contemporary cognitive science: embodiment and distributed cognition. First, it questions the supposition that the body is a non-problematic, natural, self-standing phenomenon, commonly assumed in the literature on embodiment. It, also, defines mechanisms of distributed cognition in terms of the multimodal semiotic actions (the multimodal actions not only represent cognitive processes, they participate in, and structure such processes), and it provides evidence for its principles. While MRI practitioners localize cognitive functions in the brain of single individuals, the analysis of their practice reveals that cognition is always relative to the cultural and social world of embodied subjects.

CHAPTER 1

INTRODUCTION

1. INTRODUCTION

1.1 WHAT IS FMRI?

A significant force of rhetoric has been brought to bear on the idea that we live in an increasingly visual world:

> Modern life takes place onscreen. … Human experience is now more visual and visualized than ever before from the satellite picture to medical images of the interior of the human body. … In this swirl of imagery, seeing is much more than believing. It is not just a part of everyday life, it is everyday life" (Mirzoeff, 1999: 1.)

One of the crucial components of this *visual world* is scientific visualization. With technological advances, we are witnessing an increasing prominence of visual evidence in scientific work (e.g., MacEachren, 1995).

A key modern digital imaging system used for scientific and medical purposes is Magnetic Resonance Imaging (MRI) technology. The goal of MRI is to provide detailed static images of the anatomical structure of internal body parts, such as 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 brain tissues are induced to emit a signal that is detected by the computer, where the signals, represented as numerical data, are converted into visual representations of the brain.

In addition to the production of static images of interior regions of the body, 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 during that task. Thanks to this technique, scientists have access to static images or *maps of active processes within the brain*.

1.2. DISCOURSE ON FMRI

My research project on fMRI began in the summer of 2002 when the University of California, San Diego (UCSD) inaugurated its $13.5 million *Center for fMRI*. The center was announced as *the largest brain imaging facility dedicated to research in the Western United States* (UCSD News, October 30, 2000). Here are two short but representative quotes from speeches delivered at the Center's opening ceremony:

Edward W. Holmes – Vice Chancellor for Heath Sciences / Dean, School of Medicine

> (…) With this dedication we announce the availability of this powerful imaging facility that will serve a wide variety of investigators: neurologists, psychiatrists, cognitive scientists, radiologists, engineers, biologists, and chemists. This state of the art resource will accelerate the pace of discovery in the studies of the brain and its function. Armed with this technology we can address the fundamental questions about what is arguably the most fascinating organ in the body – our minds.

Roderic Pettigrew – Director, National Institute of Biomedical Imaging and Bioengineering

> We can now visualize the mind in imaging the brain. With this kind of resolution we are now able to attach the intangibles to tangibles to see the mind in the brain. Memory, forgetfulness, truth, lie, happiness, sadness can be visualized and thereby studied.

The quotes testify to the pervasiveness of two themes in the discourse on fMRI:

1) The collaborative nature of the enterprise.
2) The centrality of fMRI technology for disclosing the workings of the human brain/mind.

Because of its complexities and its high cost, an adequate use and understanding of fMRI requires extensive collaborations at various levels. Speakers at the UCSD's fMRI Center's opening ceremony frequently mentioned collaboration between different *actors*: institutions, research fields, and individuals.

The extensive effort that various levels of collaborations require finds its justification in the particular significance of the endeavour: after more than 2000 years of speculation in Western thought, recent explorations of the brain, its anatomy and function, promise a scientific access to the enigmas of the human mind[1]. It is believed that such an access is attainable through advancements in imaging technology: images are mediators between the mind and the material and physical reality (i.e., the brain). In other words, the discourse on fMRI assumes that once visualized, *invisible* mental phenomena – rather than just brain states - can acquire status of empirical facts. The idea of understanding mental phenomena through visualization of functions is implied by the conceptual compression between the brain and the mind whose understanding is attainable through the technology, as proposed in Holmes' quote, and is clearly expressed by Pettigrew where he links the images' resolution, visibility, and tangibility with mental phenomena such as memory, forgetfulness, truth, lie, happiness, and sadness.

In fact, the quote from Pettigrew's speech highlights one of the central features of the fMRI evidence - its spatial character. While other neuroimaging techniques such as electrophysiological recordings (i.e. electroencephalography (EEG) and magnetoenchephalography (MEG)), provide exact information about the temporal dynamics of electophysiological activation and deactivation processes occurring in the millisecond range, metabolic imaging techniques such as PET and fMRI are extremely powerful in localizing

---

[1] The idea of a *scientific* access to the mind greatly determines the criteria according to which certain research fields and disciplines (natural sciences and engineering, rather than social sciences and humanities) are elected as eligible participants in the collaborative effort of the study of the mind.

brain structures that are active during a cognitive task, while providing only a rough picture of the temporal dynamics of brain processes (Pulvermüller, 1999). In other words, the strength of fMRI maps is their aptitude to indicate where in the brain cognitive processes like memory, attention, language, spatial cognition, etc., take place (1-1).

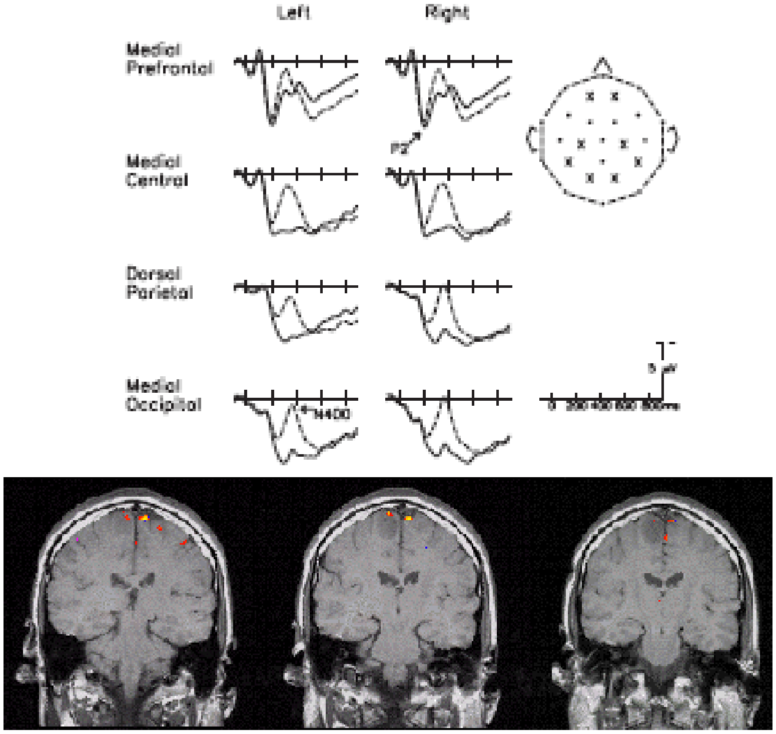


FIGURE 1-1: Representation of electrophysiological recordings (top) vs. fMRI

One such process that can be studied through the use of fMRI maps is vision (e.g., Sereno et al., 1995, 2001; Tootell et al., 1998; Wendell, 1999). By identifying location and structural organization of visual brain areas, scientists extensively work on understanding our *visual world* and the *visual nature of our experience*. But while such a research is of extreme importance, a full understanding of visual phenomena requires an effort that, in addition to the knowledge of neuroscience, would also take into consideration an understanding of socio-cultural environment of practice. In order to shed light on such an environment, the aim of my project is to situate brain imaging studies in their every-day context of practice. This allows me to focus on socio-cultural and embodied aspect of fMRI work, and question the widely spread assumption that our understanding of images and the reality they represent is achieved exclusively through an “abstract sense of vision” (Lenoir, 2004: XXII).

2. BACKGROUND

2.1. STUDYING SOCIO-CULTUAL ASPECTS OF BRAIN MAPPING PRACTICE

To investigate socio-cultural dimensions of brain imaging is not unique. In the following portion of the text I will briefly review the research that is being done in the domain.

Joseph Dumit (2004), a social scientist studying brain mapping, combines science and technology studies (STS) with cultural anthropology and history to tackle two related problems. On one hand, he explores how positron emission tomography (PET) images serve different agendas and acquire different meanings. By investigating how images “travel” – e.g., from a science laboratory, to scientific publications, to popular media, and back - he focuses on the processes of production, translations, use, and appropriation of facts.

In addition, Dumit explores how brain images participate in processes of categorizations, and identification. To do so, he investigates how relationships between images and personhood get established. Dumit points out that PET images are constructed and framed to allow for each person to be subsumed under a certain type of brain, scan, and diagnosis: “These brain images make claims on us because they portray *kinds* of brains. As people with, obviously, one *or* another kind of brain, we are placed among the categories that the set of images offers” (2004: 5).

Simon Cohn, a medical anthropologist, conducts a very closely related research. He investigates how visual representations are used within science and how they are re-interpreted by non-experts. He does so by studying how brain imaging technology affects the traditional cultural boundary between mental and physical illness. Cohn inquires if the visualisation technology produces consequences for the understanding people have of the mind and body, normality and illness (http://www.strangebedfellows.net/simon.htm).

An example of such an issue, and of how problems of categorizations and definition of mental illness appear in discourse on fMRI, can be found in Judd's speech delivered at the UCSD's fMRI Center's opening ceremony. Judd talked about the influence that brain images have on establishment of new policies and understanding of mental illness: "… our ability to demonstrate this - a picture is worth thousands words in this case- has been very telling in Congress to begin to change policies to bring mental illness up to parity with physical diseases…". Brain images allow scientists to define certain brains as representative of a mental illness and thus anchor "mental illness" in its biological base: "We were able to irrefutably and unequivocally establish that mental disorders are brain based biological diseases and that they are not characterological flaws and they are not moral weaknesses or any other mythologies that had been attributed to the mental illnesses".

Another researcher interested in social aspects of fMRI is Anne Beaulieu. Beaulieu examines questions of visualization and digitalization of knowledge in neuroscience. While working on her doctoral dissertation, she interviewed fMRI researchers to find out what role brain images play in the production of knowledge (2002). Her data show that scientists who use fMRI and PET reject the importance of brain representations as visual data. Beaulieu argues that because it has been regarded as appealing first to the senses, as opposed to reason, and hence is seen as lacking a solid relationship to the truth, visual evidence is judged as not having a particularly high position in the hierarchy of types of scientific evidence. But while Beaulieu's research successfully explains some aspects of the discourse on fMRI, because it is based on what researchers report on their practice, it leaves unexplained how scientists in fact engage with visual data during their research practice.

Beaulieu's recent article (2004) is mainly concerned with digital information and the networked character of research in neuroscience. By studying brain databases (i.e., collections of brain scans), she points out consequences generated by the translation of an object, such as a biological and material brain, into digital information. Manipulability of the virtual brain brings directly into the play questions of individuality and standardization, as well as the status of the object itself:

> As an informational object, the virtual brain lends itself to different sets of manipulations that extend, refine, and dissolve its individual component. The original scan is both lost and extended in this series of transformations. Collections of virtual brains therefore seem to have a particular combination of functions, as data, experimental system, and model. Researchers grapple with this hybrid status of their atlases, and are often at a loss to name this beast, which is at once an atlas, a database, a summary of data and an object for future research" (2004: 379).

While research, largely grounded in analysis of discourse on fMRI, such as Dumit's, Cohn's, and Beaulieu's work, successfully engages with questions concerning objectivity, standardization and generalization, it remains somewhat removed from particular circumstances of scientific practice, and thus leaves questions of materiality and embodiment fairly unexplored. For example, Beaulieu (2004), when characterizing the nature of the new types of digital, hybrid objects, such as brains converted into "images of numbers" (Tufte, 1983), does not inquire if such digital objects and their understanding conserve some aspects of their 3D materiality. The question then is if fMRI practitioners, in order to work with and understand experimental data, translate *digital brains* into conceptual entities that call for some elements of materiality. In the present project my intention is to stay attentive to discursive constructions, while examining how embodied humans interact with the material conditions in which they act (Hayles, 1999: 195).

Problems of objectivity, standardization and generalization are also addressed by Andreas Roepstorff. While seeing knowledge as an empirical problem to be examined, he asks: "What is it that a certain group considers knowledge? Who would actually take this to be true and/or useful? How do they go about establishing this as knowledge? (Roepstorff, 2004: 1106).

But in addition to his research of the discourse on fMRI, Roepstorff studies the practice of fMRI. Particularly interesting is his examination of the interaction between the experimenter and the experimental participant (Roepstorff, 2001, 2003; Jack & Roepstorff, 2002; Roepstorff & Frith, 2004). Roepstorff argues that such an interaction needs to be factored into an understanding of brain imaging experimental results. Cognitive scientists and neuroscientists understand cognitive function in terms of top-down or bottom-up processes: action can be controlled automatically by sensory inputs, or governed by conscious internal knowledge that the organism has on its own. However, their distinction leaves out the fact that when a human participant is involved in a brain imaging experiment, his/her cognitive actions are not channelled only by bottom-up or top-down processes. Instead, they are also significantly *guided by the experimenter* (Roepstorff & Frith, 2004). This perspective points out the social and cultural character of cognitive processes:

> Performance on the task is a joint enterprise, but it is only rarely described as such in the cognitive literature. Outside the laboratory this strange distinction between participant and experimenter no longer holds and the shared nature of our endeavours is much more obvious. Once the analytical attention is shifted from the actual performance of the participant during the experiment to the 'setting up' of the experiment, the inherently shared aspects of the cognitive experiment become very apparent (ibid: 195).

To explain the direct involvement of the external factors in the cognitive processing of an experimental participant, Roepstorff & Frith (ibid.) invoke the notion of "shared scripts for action". Every experimental session starts with an experimenter giving a set of instructions to the experimental participants. The result of such a communication is the extablishment of "common, shared representations about the nature of the task" (ibid: 192). During the experiment, the participant enacts a script *acquired from the experimenter*. Roepstorff & Frith point out that the origin of the control of action is "outside the brain of the participant, namely in the mind of the experimenter" (ibid: 192). Scripts are "imposed" (194) or "directly received" (193) *from the experimenter*. The experimenter is in turn "controlled by scientific knowledge provided by his predecessors and peers" (194).

This explanation of cognition and the questioning of brain imaging experimentation are of an unquestionable significance. It is rather rare for a researcher to question assumptions about human cognition while studying actual experimental procedures in cognitive science. Yet, Roepstorff & Frith’s story is still a disembodied and abstract account of cognition. Their explanation does not seriously consider how processes such as “scripts sharing”, for example, take place. In their account, the cognition of the experimental participant is reallocated from his/her “head” into the “head” (e.g., scripts and internal representations) of the experimenter, and eventually into the “heads” of her/his predecessors and peers.

The present study, while revealing social and cultural aspects of cognitive processes such as meaning-making, seeing or learning, aims to question the widespread view of cognition based on the idea of rational, cognitive agents. Instead, it investigates cognition as an embodied process, spread across a social and material world of action, rather than a structure encapsulated in the heads of individual agents. The question then is: how do body and socio-cultural world of practice, in addition to the mental world of an individual (e.g., experimental participant or experimenter), become involved and constitute meaning-making and action for MRI scientists?

2.2. STUDYING SCIENCE AS A PROCESS OF DISTRIBUTED COGNITION

The point of departure for the present study was an interest in the collaboration among scientists and the multidisciplinarity of the new fMRI center, i.e., the first theme identified in the introductory section. This direction for research represents a challenge: How can somebody interested in meaning-making and learning investigate multidisciplinarity without reverting to a study of individual scientists as cognitive agents? In order to avoid such a traditional perspective, my plan was to regard the fMRI center, rather than single scientists, as a cognitive system that produces scientific knowledge. The plan parallels a larger research trend.

Since 1980s, researchers started to see cognition as socially and culturally embedded, situated, distributed between people, and rooted not only in mental processes and computational inferences, but in an externally given and real world of physical objects, artefacts, and cultural practices (Engestrom, & Middleton, 1996; Lave, 1988; Middleton & Edwards, 1990; Neisser, 1982; Norman, 1988; Rogoff & Lave, 1984; Suchman, 1987). One of the most important texts in the field, and the specific theoretical foundation for my project, is Ed Hutchins' *Cognition in the Wild* (1995).

Hutchins' *Distributed Cognition* (DCog) approach is concerned with cognition as an activity undertaken in social and material settings, rather than a solitary and internal process of a single thinker. DCog treats cognitive processes as distributed across the members of a social group as well as the structure in their environment (Hollan et al., 2000). Focus is on cognition as produced at the level of a socio-cultural system, such as the fMRI center. One can, for example, investigate cognition as a process distributed across fMRI scientists, technical equipment of the center, the powerful MRI magnets, computers, and brain images.

The change in perspective from an individual thinker to a global system is important for a straightforward reason: a study of cognitive processes of a single individual does not always allow one to predict the properties of the system. A distributed cognitive system may have emergent characteristics that are generated by interactions among properties that are not present in any individual element of the system.

In order to describe such external mental processes, Hutchins (ibid.) uses classical cognitive science terminology. He defines culturally and socially distributed processes by adopting the principal metaphor of cognitive science – *cognition as computation* – previously used only when internal mental processes were described. Cognition thus is defined in terms of *computation as the propagation of representational state across a variety of media* (Hutchins, ibid.: 49). Hutchins claims that by defining cognition in such a way no special commitment must be made as to the nature of the computations that are going on within individuals. Nevertheless, DCOg doesn't simply assume a classical stand toward internal mental representations. Internal cognitive processes are always part of a larger computational system: "(A)lthough some of the representations are internal, they are still all cultural in the sense that they are the residua of a process enacted by a community of practice rather than idiosyncratic inventions of their individual users" (ibid.: 130). Similarly, Hutchins describes external mental representations by using computational vocabulary while making clear that this type of description need not be understood in strictly traditional terms:

> I propose a broader notion of cognition because I want to preserve a concept of cognition as computation, and I want the sort of computation that cognition is to be as applicable to events that involve the interaction of humans with artefacts and with other humans as it is to events that are entirely internal to in individual persons. …. The actual implementation of many interesting computations is achieved by other than symbolic means. For our purposes, 'computation' will be taken, in a broad sense, to refer to the propagation of representational state across representational media. This definition encompasses what we think of as prototypical computations (such as arithmetic operations), as well as a range of other phenomena which I contend are fundamentally computational but which are not covered by narrow view of computation (118).

This broad view of cognition and computation produces a particularly valuable outcome for studies of real-world practices: cognition can now be directly observed. According to the DCog position, "… while some underlying processes (inside people's heads) remain obscured, a great deal of the internal organization and operation of the system is directly observable". (ibid: 129). Not only co-participants in the activity, but also ethnographers can observe cognition as it unfolds through practice.

The distributed approach to cognition appears particularly suitable for a study of science where knowledge is produced through interaction between scientists and technology. Yet, the research in DCog, despite its suitability, hasn't been directly applied to scientific practices. Instead, such phenomena have been analyzed at length by proponents of *laboratory studies*. Laboratory studies - among which the most well-known are Karin Knorr-Cetina's (1981) study of a protein chemistry laboratory in Berkeley; Bruno Latour and Steve Woolgar's (1979) study of a biochemistry laboratory at the Salk Institute; and Michael Lynch's (1985) study of a neuroscience laboratory at University of California, Irvine - describe scientific practice in its "natural" setting (Lynch, 1985: xiii). By doing so, their research stands in contrast to the traditional trend in science studies where science was studied as a phenomenon divorced of any specific practical instance of scientific practice:

> *Science* is portrayed as a general phenomenon in science studies. This is to say that science studies depict a science that is trans-situational, and not a matter of, for example, watching scientists at work to see what they are doing in any immediate way. Science studies are replete with systems of explanation, rules of logic and method, principles of theory construction and experimental design, and norms of science which are formulated so as to cover all variants (or some disciplinary sub-unit) of *Science*. These rules of evidence, norms of science, etc., cover a science which is more *extensive* than any instance of its performance. Regularities which are attributed to science are not of the manner of empirically situated "facts" observed in the behavior of scientists, but instead describe persistent themes which are most congruent with communally held "versions" of how science is or should be (Lynch, 1985: 144).

The orientation toward practices undertaken in natural settings is akin to the DCog perspective. Giere (in press) points out that the work of Bruno Latour (1986) and Karin Knorr-Cetina (1999), even though it does not explicitly evoke DCog framework, implicitly applies its principles. Knorr-Cetina's analysis of the High Energy Physics Laboratory (at the European Center for Nuclear Research) suggests an understanding of laboratory work in terms of a cognitive system. In such a system, cognition is distributed across a large number of scientists, technicians, and other support personnel. By analyzing such a system, Knorr-Cetina claims that one cannot identify any individual agent that produces knowledge in the system. Knowledge is always generated at the level of the system. Rather than attributing the role of an *epistemic subject* to a single individual, Knorr-Cetina attributes such a role to the extended experiment itself. Latour (1986), on the other hand, looks at the history of science and technology to point out how cognitive processes of scientific problem solving consist of interactions between a person and physical, external representations. Scientific inventions depend directly on the development of external representations as well as the instruments that produce such representations. Cognition, thus, is distributed across a network that contains various *actants* (both humans and nonhumans), and is a locus of transformations of external representations.

Significant for my project is Giere's clarification that the DCog framework does not only concern global *distributed cognitive systems* (like the nuclear physics laboratory analyzed by Knorr-Cetina) where cognition is spread across a large number of individuals that engage with the material world. The framework also pertains to *locally distributed cognition* (exemplified by the instances of scientific practice described by Latour, and clearly present in Andy Clark's work on embodiment, 1987, 1997, 2001).

> Andy Clark's (1997) spirited advocacy of cognition as being distributed focuses on what I would call *locally distributed cognition*. A person working with a computer would be a paradigm example. The person with a computer can perform cognitive tasks, such as complex numerical calculations, that a person alone could not possibly accomplish as accurately or as fast, if at all. When it comes to arguing for extending the concept of *mind* beyond the confines of a human body, he invokes a more primitive example, that of a man with a defective memory who always keeps on hand a notebook in which he records all sorts of things that most of us would just remember (1997, 213-18). Clark claims that the person's mind should be thought of as including the notebook. For this person, the process of remembering something important typically involves consulting his notebook. The notebook is part of his memory. A major component of Clark's argument for this position is that for someone else deliberately to damage the notebook would be equivalent to assaulting the person. The notebook is as crucial to this man's normal cognitive functioning as is the left part of his brain (Giere, ibid).

While for Giere locally distributed cognition concerns interaction between an individual and external representations or artifacts, the term can also be used to cover situations where one or various agents use their bodies as part of the everyday practical dealings. We can, for example, observe an individual and analyze how meaning-making, interpreting and learning is distributed across different semiotic modalities, such as gesture, body orientation, talk, gaze, etc. that participate in the culturally meaningful and socially organized practice. The focus is then on *cognitive artifacts* on the one hand, and *multimodality* on the other as part of human action and interaction.

A variety of external structures can function as *cognitive artifacts* (Norman, 1991; Hutchins, 1999): computers, geographic maps, or stellar constellations. By coordinating functional skills with these structures, produced through human practices or present in nature and opportunistically used, cognition can be aided, or improved (ibid.). However, cognitive artifacts are not only an aid for cognition; they are part of distributed cognitive processes. In this way, external resources, such as the fMRI scanner, function as "a means of expanding and reshaping the space of human reason" (Clark 2001).

In addition to the material world, the human body takes on a central role in the study of distributed cognition. In accordance with the embodied view (Clark, 1987, 1997), material aspects of agents' bodies are highly significant for any adequate characterization of sense-making and reasoning. Meaning and action are negotiated and coordinated through linguistic means, as well as through body orientation and gesture. Hands, eyes, position of the head and position of the body are all available for manipulation in action space (e.g., Hutchins & Palen, 1993; Goodwin, 1996: 2000a). Consequently, the organization of the mind is conceived as an emergent property of interaction among different communicative modalities and external resources employed by one or more practitioners (Hutchins & Palen, ibid.).

To see how such aspects of sense-making and knowledge sharing are developed through situated embodied practices, the analysis has to be scaled down to a micro level of multimodal interaction. This level of analysis hasn't been commonly realized in studies of science and technology. While the research by Knorr-Cetina and Latour, for example, can be characterized as applying principles of DCog (both global and local, respectively), we still know very little about how multimodality and embodiment play a role in the practice of science. What does actually happen in those moments in which scientists engage their bodies in the process of sense-making in the disciplinary specific *Lebenswelt* ?

2.3. MULTIMODAL INTERACTION IN SCIENTIFIC PRACTICE AND CONCEPTUAL INTEGRATION

A researcher who has studied real-world scientific activities while developing a new theory of multimodality is Charles Goodwin. Goodwin analyzes practices situated in diverse settings: e.g., court room procedures (e.g., 1994); playground activities (e.g., 2002); interaction with an aphasic man (e.g., 2000b; 2003a; 2003b); as well as scientific activities such as archaeological excavations (e.g., 2003b); oceanographic research vessel (1995); and a team of geochemists (1997). In his writings Goodwin claims that cognition should not be divorced from socially organized, situated activity accomplished by members of a specific *community of practice* (e.g., 1994) in a relevant setting. Therefore, cognitive processes do not only pertain to internal mental representations, but comprise material objects, various embodied and social structures and processes. The integration and organization of such diverse phenomena is provided by the *situated activity system* (e.g., 1997). By showing how cognitive processes and meaning-making are always lodged within endogenous communities of practice, while being constructed within a socially constituted world of relevant activities, rather than being a reflection of mental processes of an isolated actor (1997:116), Goodwin's work is very akin to the DCog perspective. Yet, Goodwin's study adds to such a perspective a strong focus on the details of interaction developed through coordination among various semiotic modalities (Goodwin, 2000c).

In order to study human cognition and action, Goodwin videotapes situations where multiple actors align with each other while carrying out an activity. He suggests that

> … a primordial site for the analysis of human language, cognition, and action consists of a situation in which multiple participants are attempting to carry out courses of action in concert with each other through talk while attending to both the larger activities that their current actions are embedded within, and relevant phenomena in their surroundings (2000c: 1492).

The videotaped data allow Goodwin not to be bounded only to language, or to any one semiotic modality, but to analyze coordination and mutual elaboration among various semiotic resources and the environment where the action takes place. One of the distinct features of Goodwin's work is the meticulous detail with which he explores how actors constitute their social world though juxtaposition among language, body orientation, gesture, and material structure in the environment (e.g., 1994, 2000c). This type of analysis, which looks at coordination between different semiotic modalities, highlights similarities between everyday, generic human activities, and specialized, scientific practices:

> … with this framework it is … possible to demonstrate how activities that may initially appear to be quite unique and exoteric, such as the details of scientific work, are in fact built through use of far more pervasive, indeed generic practices for the accomplishment of action within situated human interaction (2000c: 1506).

A particularly appealing aspect of Goodwin's work is his exploration of the embodied apprenticeship. Goodwin points out that in order to become competent in the use of a professional category, a novice, learns through situated activity, rather than acquiring a set of pre-established, abstract and universal categories. Such learning consists of an active involvement of gesture, talk, gaze, body orientation, and the historically built material world.

By investigating its public visibility, Goodwin describes how the body functions as an interactively organized locus of the production and display of meaning and action. While it is used to display mutual orientation, it is also involved in the production of gesture. In respect to gestures, Goodwin's analysis points out that they do not simply externalize internal mental representations (McNeill, 1995, 2000). Instead, gestures emerge from, and point to the embodied activity through which they become visible (Goodwin, 1997: 128).

To further develop Goodwin's line of research as a part of *practice-based theory of knowledge and action* (see Goodwin, 1994: 606) the present work borrows some insights from *Conceptual Integration Theory* (Fauconnier & Turner, 2003; Coulson, 2000) (or *Blending Theory*). Similar to the well known *Conceptual Metaphor Theory* (Lakoff, 1987), Conceptual Integration Theory unifies the analysis of metaphor with the analysis of other linguistic and conceptual structures. The theory is involved in identifying systematic projections of language, imagery, and inferential form. A significant aspect of both theories is a focus on *embodiment*. According to the embodied view of cognition and semantics, conceptual forms emerge from the structure of the human body. For example, Lakoff (1987) explains that the understanding of technically sophisticated aspects of science is often based upon human biological capacities and the experience of functioning in a physical and social environment. In this way one can understand abstract aspects of mathematical procedures in terms of physical actions (Fauconnier & Turner, 2003 speak about *concepts being on human scale*).

What makes Blending Theory unique is its focus on novel and often short-lived conceptualizations. Instead of accounting for stable knowledge structures represented in long-term memory, the theory models the dynamic evolution of speakers' on-line representations

(Coulson & Oakley, 2000). To account for such phenomena, Fauconnier and Turner (1998) conceive of a set of general cognitive processes – *conceptual integration or blending* - used to combine conceptual structure, and posit *mental spaces* (Fauconnier, 1994), on which conceptual integration operates. Mental spaces are short-term mental scenarios constructed on-line. Blending takes place in a *conceptual integration network* (Figure 1-2), an array of mental spaces which typically includes two input spaces, a generic space, and a blended space. *Input* spaces represent information from discrete cognitive domains, a *generic* space contains structure common to the two inputs, and the *blended* space contains structure from both inputs, as well as its own emergent structure. Emergent structure arises out of the imaginative processes of blending. The first process is called *composition*, and involves the juxtaposition of information from different spaces; *completion* occurs when part of a cognitive model is activated and results in the activation of the rest of the frame. Finally, *elaboration* is an extended version of completion that results from mental simulation, or various sorts of interaction with the world as construed with blended concepts. Hence, blending is an on-line, real-time process that creates new meaning through the juxtaposition of familiar material.

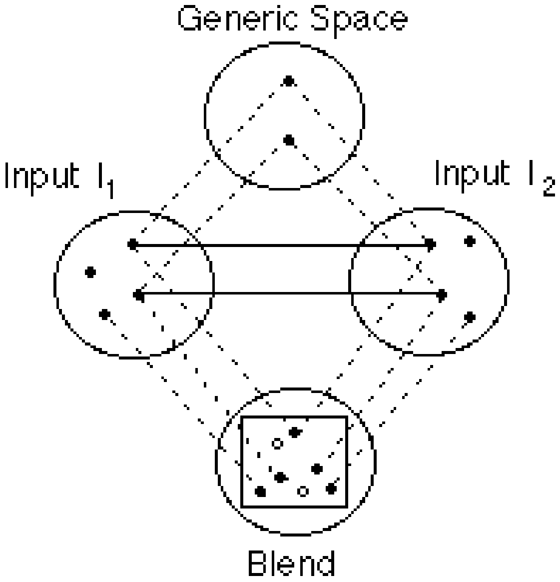


FIGURE 1-2: Conceptual Integration Network

One of particularly interesting aspects of this semantic theory is its focus on imagination and creativity. Fauconnier & Turner (ibid.) suggest that principles of conceptual integration depend on the human imaginative capacity. The imaginative capacity operates in every instance of meaning construction. While formal approaches considered such phenomena to be unanalyzable, blending theory posits as its principal goal a detailed study of processes of imagination. In the introduction to their book *The Way We Think*, Fauconnier and Turner write:

> The theme of this book is what the form approaches have taken for granted, assumed as given: the operations of imagination, integration, and identity. These operations - basic, mysterious, powerful, complex, and mostly unconscious - are the heart of even the simplest possible meanings. We will show that they are the key to the invention of meaning, and that the value of even the simplest forms lies in the complex emergent dynamics they trigger in the imaginative mind. We will argue that these basic operations are more generally the key to both everyday meaning and exceptional human creativity. Surprisingly, but as it turns out crucially, even the most basic forms, the chestnuts of the form approaches, are prompts for massive imaginative integration (ibid:6).

One of the challenges for the present study is to describe processes of imagination as practical aspects of scientific work. By adopting some of the insights from conceptual integration theory, I point out how practical ways of problem solving in MRI laboratory, rather than being epistemologically “special”, share some common principles with other non-scientific, every-day instances of reasoning and meaning construction.

Yet, how can one reconcile an analysis of reasoning and sense-making, usually concerned with internal mental processes, with a study of social and public practice? While embracing an embodied view of cognition, theories from cognitive semantics, blending theory included, focus primarily on cognitive processes of individuals (Goodwin, 2000a):

> … Much study of metaphor has taken as its point of departure the embodied *experience* of the speaker, e.g., the way in which metaphor emerges from the structure of the human body, its position in a world structured by phenomena such as gravity (e.g. the pervasive relevance of *Up* and *Down* in human cognition and language) and ‘preconceptual structure of experience’ (Johnson, 1987: 15). While providing valuable insight into many kinds of conceptual organization, such focus on

> the interior life of a single actor does not develop a systematic framework for investigating the public visibility of the body as a dynamically unfolding, interactively organized locus for the production and display of relevant meaning and action (Ibid.: 1517).

Though this critique has to be seriously taken into account, the general orientation toward conceptual processes of individuals present in the blending theory does not preclude an enlargement or dislocation of the perspective to the public phenomena of action and interaction. A first attempt to move the theory in this direction is the idea of *material anchors for conceptual blend* proposed by Edwin Hutchins (2005).

Hutchins (ibid.) pointed out the existence of conceptual integration networks where one input space is a conceptual space in the usual sense, while the other one has the structure of a physical object or event. The examples of such structures are watch, method of loci, or people standing in line. Hutchins' examples show how conceptual relations can be mapped onto relations among the material elements. One of the emergent properties of such constructs is the stability of conceptual representation. Yet, in this kind of networks, the products of integration between spaces are still largely accomplished internally.

The question that the present work poses is: Can this type of analysis be further expanded to talk about a process of integration that includes more than one input space expressed in material form, or even articulated through an embodied semiotic enactment? Furthermore, can the process of integration and its final products be first of all and largely produced in the interpsychological world of everyday practice?

To tackle these questions I describe cases in which structures that exist in the external world of practice are brought into coordination in a way that suggests a process of their integration. To see processes of integration performed in socially shared material spaces of practice I turn to Charles Goodwin's concept of *semiotic fields* (2000). My analysis focuses not only on stable representational forms and their physical manipulations, but also on integration among material and embodied forms through engagement of dynamic semiotic modalities.

3. FMRI IMAGES

Short after I started observing practices at the new fMRI center, as well as activities that take place in fMRI laboratories at UCSD campus, I have become aware that the question of collaboration (e.g., the first topic identified at the fMRI center's opening ceremony, and my initial topic of interest) is inseparable from the question concerning brain images (e.g., the second topic identified at the opening ceremony). As the fMRI center and the fMRI laboratories are conceived as socio-material networks consisting of human as well as nonhuman elements that generates knowledge, collaborations and interactions among parts of the networks are their driving force. A large number of such interactions take place around brain images, or - brain images become the locus of such interactions. Scientists engage with the experimental data and communicate among themselves by using images; images are the basis for the production of other images, while also being the published results of the laboratory work. To study collaboration I had to first seriously take into account questions concerning images and the sense-making processes in which they participate.

As the discourse from the opening ceremony reveals, with the development of brain imaging techniques, images have become central for *a scientific understanding of the human mind*. Brain images do not only show anatomical structures of living tissues, invisible to our unaided eye; they *visualize biological processes* that are not in themselves visual in order to *allow* scientists to form claims about the nature of human cognitive processes. Such images are produced through socio-cultural practice, and have important social and cultural consequences. Social scientists, as mentioned above (Beaulieu, ibid., Cohn, ibid. Dumit, ibid., Roepstroff, ibid.), have researched the role of images within the context of the debate about the human mind, while emphasizing the social and cultural role that such images play. They point out that brain images, in order to become meaningful, always involve cultural and social processes of interpretation:

> For better or worse, depending on the perspectives one holds and the research bias one has, images are the raw material of scanning technologies like MRIs. In other words, the brain is visualized at a topological level, mapped according to various levels of excitation of a chemical and electrical nature, and researched and treated through the knowledge gained. MRI technology captures the molecular structure of human tissue, which produces enough of a magnetic charge to allow the signals to be reassembled into images. This is primarily biological model and leaves many questions unanswered about mind, thought, and relationships between perception and thinking. In particular, the issues of how images are used to explain biological processes needs to be framed by cultural argument and cultural criticism.
>
> These lacunae would not be an issue except that the use of images entails far more than the transparent relationship of scanning to results would suggest (Burnett, 2004: XV-XVI).

As Burnett (ibid.) points out, activities of *use* and *interpretation,* which take place between the brain scanning and formation of scientific results, are still an undiscovered territory. The discovery of such a territory requires a particular kind of socio-cultural and phenomenological (I would claim) analysis. We still don't know how the processes of interpretation happen at the level of actual, situated, embodied practice. What happens when scientists analyze the experimental data by directly working with images? How do they talk to each other while referring to images? How do they teach novices to read images? How do they select images for publication, and negotiate the meaning of published images with the larger scientific community? In order to start confronting these questions I had to look at moments and places of fMRI practice where brain images cluster and interact with multiple material and embodied socially organized structures.

As repeatedly pointed out by scholars in social studies of science (e.g., Kuhn, 1962; Mindell, 2002; Pickering, 1995) the material world of tools and instruments is a crucial aspect of every scientific enterprise. Accordingly, the meaning of fMRI images is relative to their interactions with other material elements ranging from extremely complex instruments such as fMRI magnets to paper, pencils, and blackboards. Such structures always reveal their *multivoicedness* (Bakhtin, 1986; . Cole, 1998; Lotman, 1990; Wertsch, 1998) and chains of *connotations* (Barthes, 1988; Eco, 1976) in which they participate. The relationship between MRI images and other artifacts that populate laboratory setting is frequently mediated through various more or less stable semiotic forms. Scientists continually coordinate brain images with other *immutable mobiles* (Latour, 1986), like charts, graphs, brain atlases, etc. (Chapter 4) to make sense of and strengthen their experimental data.

In addition to instruments, tools, and other visual representations, the present work explores the intersection between brain images and various multimodal and dynamic semiotic resources. Neuroscientists, while engaging with brain images as well as with their interlocutors, coordinate such images with speech, gesture, body orientation, and gaze. The active inclusion of such modalities in the processes of image interpretation emphasizes its embodied character.

By looking at the coordination and mutual dependence among brain images, instruments, tools, other visual representations, as well as multiple semiotic modalities, I describe how multimodal interaction between neuroscientists and the material environment of practice is sensibly related to the contextual production of fMRI images. The transitory semiotic constructs reveal how embodied and situated interpretations of brain images feature in the course of laboratory work. To generate images as *representations of human cognition*, neuroscientists engage in knowledge production processes as lived work of MRI imaging.

To look at the embodied aspects of everyday meaning-making activity, the focus had to be scaled to the multimodal semiotic enactments of visuospatial representations. This move entailed a shrinking of the original unit of analysis: with the proceeding of the study, my focus moved from the fMRI center to a much smaller unit – everyday scientific practices conducted in three fMRI laboratories.

Does this focus on the micro-level of practice necessitate a loss of the “big picture”? Is such an analysis inadequate for illuminating social and cultural context of research in neuroscience? No. Rather than stepping outside of the phenomenal field of the practice to reveal the “big picture”, I adopted a prospective akin to *ethnomethodology* (Garfinkel, 1967) whose goal is “to insist that specifications of such contexts are invariably bound to a local contexture of relevancies” (Lynch, 1993: 125).

4. ETHNOGRAPHY AND MICRO ANALYSIS OF MULTIMODAL INTERACTION OF SCIENTIFIC PRACTICE

In order to examine processes of meaning-making in multimodal interactional complexes, I combine traditional long-term participant observation with the micro-analysis of specific occurrences of events and practices. The long term participant observation, or ‘ethnography’, is a study that stems from researcher’s immersion in the group’s natural environment and practices. The aim of ethnography is directed toward identifying and explaining complex social structures within the group or culture studied. An ethnographer needs to immerse herself and participate in a group or culture, while, at the same time, trying not to be manipulative of group behaviors. Thus, *doing ethnography* requires more than following a set of procedural steps:

> In understanding what ethnography is, or more exactly *what doing ethnography is,* that a start can be made toward grasping what anthropological analysis amounts to as a form of knowledge. This, it must immediately be said, is not a matter of methods. From one point of view, that of the textbook, doing ethnography is establishing rapport, selecting informants, transcribing texts, taking genealogies, mapping fields, keeping a diary, and so on. But it is not these things, techniques and received procedures, that define the enterprise. What defines it is the kind of intellectual effort it is: an elaborate venture in, to borrow a notion from Gilbert Ryle, ‘thick description’.… the point is that between what Ryle calls the "thin description" of what the rehearser (parodist, winker, twitcher …) is doing (‘rapidly contracting his right eyelids’) and the "thick description" of what he is doing (‘practicing a burlesque of a friend faking a wink to deceive an innocent into thinking a conspiracy is in motion’) lies the object of ethnography: a stratified hierarchy of meaningful structures in terms of which twitches, winks, fake-winks, parodies, rehearsals of parodies are produced, perceived, and interpreted, and without which they would not (not even the zero-form twitches, which, as a cultural category, are as much nonwinks as winks are nontwitches) in fact exist, no matter what anyone did or didn't do with his eyelids (Geertz, 1973: 5-6).

Unlike traditional ethnography, that primarily focuses on the interdependence of group behaviors and interactions in a group, my approach aims to uncover the meaning-making processes as features of such behaviors and interactions. This is done through a micro-analysis of material, embodied, and social means of action and meaning. The two sides of the research, the micro analysis and the larger ethnography, are interdependent.

I conducted my ethnographic study in three laboratories that study cognitive processes through employment of fMRI. As a member of the cognitive science community, I entered the laboratory as a 'native' of the community while conducting an ethnographic study of laboratory practices. In this sense, I occupied the dual role of cognitive scientist and social science investigator focusing on the communication and knowledge in my research setting.

My study involved fifteen participants, and employed a variety of data collection methods, including direct observation, video recording, semi-structured interviews, and analysis of documents, such as research articles, laboratory manuals, and scientific correspondence. The use of digital technologies for the collection, storage and manipulation of data, shape the process of ethnographic research. The technology helps to make ephemeral forms of conduct – prominent features of scientific work – amenable to the repeated scrutiny of the analyst and the academic community at large (Heath & Hindmarsh, 2002).

In order to uncover how scientists construct the meaning of complex fMRI images, micro-analysis on digital video recordings (Goodwin, 2000c) of scientific practices was performed. The digital recordings were used to expose and track dynamic semiotic structures that scientists bring into coordination with the images in the process of making the images meaningful. The knowledge from my long-term ethnography was used to provide warrants for the judgments concerning how things and events are made relevant to one another in a minute instance of scientific practice.

5 RESULTS

The two primary aims of this study are:

1. To expose the way in which scientists interpret visual representations, such as digital brain images.

2. To explore communication and meaning-making as processes distributed across various embodied semiotic modalities as well as the social and material environment of practice.

These issues are examined through an ethnographic study of professional neuroscientists and their ordinary practical accomplishments in the digital and material world of fMRI laboratory.

Images matter

As pointed out in the beginning of this text, a large number of investigations have been devoted to the study of the role that visual representations play in various scientific domains (see Cartwright & Sturken, 2001; Mirzoeff, 1999). While researching the role of visual representations in brain mapping, Anne Beaulieu (2002) has claimed that neuroscientists highlight the potential of brain imaging measurement to render spatial components and anatomical referents. At the same time they downplay the visual form this information takes in order to emphasize the quantitative information it represents.

The present study - motivated by the question of how fMRI researchers interpret their experimental data in situ, rather than how they report on their research activities through after-the–fact accounts - indicates, however, that the visual form of information is central for brain mapping practice. Scientists extensively use visual representations during their laboratory work (e.g., in analyzing data or explaining the principles of the technique), as well as in their presentations of research results to wider audiences. Illustrations of this claim include the process of publishing a scientific article on brain mapping (Chapter 2), and a description of observed activities in the fMRI laboratories (Chapters 3-5). The publishing procedure shows the centrality of brain images in at least two respects. The majority of reviewers' comments were focused on visual representations (rather than written text and/or numerical data). Likewise, in order to publish the article the work of the research team largely centered on the restructuring of the originally submitted figures. In a similar fashion, data analysis practices – crucial and every-day activities in fMRI laboratories - are largely devoted to visually expressed experimental data: when scientists analyze experimental data, a significant amount of their time is spent in front of brain images displayed on computer screens.

These observations fit findings in contemporary cognitive science. Cognitive scientists emphasize the centrality of visual experience for our problem solving capacity. Since the pioneering work of Larkin & Simon (1987), cognitive scientists, and in particular students of *diagrammatic reasoning*, have pointed out that the way in which a problem is represented - whether, for example, it explicitly preserves information about the topological and geometric relations among the components of the problem - is crucial for our performance in problem solving (for reviews see Glasgow, Narayanan, Chandrasekaran, 1995; Anderson, Meyer, Olivier, 2002). Similarly, researchers have pointed out the role of *re-coding* in cognitive task performance. In re-coding an abstract mathematical problem is re-presented into a form which exploits the specific computational powers of our visual system (Kirsh, 1995)[2]. A closely related concept from *embodied cognition* (e.g., Lakoff and Johnson, 1999) is the idea of *conceptual scaling* (Fauconnier & Turner, 2002). Through processes of conceptual scaling, what is invisible or not visual (as in the case of quantitative data) is transformed into what we can directly experience (i.e., visual representations). A concept is on a *human scale* with respect to the ease with which it is produced, manipulated, or remembered.

Likewise, to understand their experimental data and results, fMRI scientists extensively use brain images. If they were limited to working with quantitative, rather than visual data, their task would be extremely difficult, if not impossible, to accomplish (see also Galison, 1997).

Interaction with images is embodied

---

[2] The idea has also been called *representational re-description* (Karmiloff-Smith, 1992) and *trading spaces* (Clark & Thornton, 1997).

The description of how scientists employ multimodal semiotic resources to engage with digital brain images in laboratory practice indicates the embodied character of the interaction. Chapters 2-5 document how scientists coordinate digital brain images not only with linguistic representations made available through speech, but also with gesture, body orientation, gaze, non-linguistic sound, and other visuo-spatial representations (e.g., other brain images, graphs, charts, etc.) to accomplish their tasks at hand. For example, in order to indicate certain aspects of brain images, scientists express such features by reaching toward the computer screen, touching it, and gesturing as if holding and manipulating round objects (Chapter 5). The involvement of the embodied semiotic modalities enable scientists to deal with, experience, and talk about their experimental data in terms of actions that take place at a level comparable to physical, real-world engagement.

What does this imply for the understanding of visual processes? Insofar as scientists, in order to interpret images, actively engage with them, we cannot talk about a *visual autonomy* (Hansen, 2004: 26). Instead, seeing is a process grounded in the body (ibid.): to notice and to share the structure of brain maps, scientists “act upon”, “feel”, and “manipulate” the information presented visually on the digital screens. Vision is an embodied process because acts of seeing are produced by the enactment of purely visual information in the world of embodied actions (Chapters 3 and 5).

Linking embodiment with socio-cultural aspects of visual representations

Yet, embodiment is typically analyzed (in cognitive science, e.g., Ballard, et al., 1997; Barsalou, 1999; Glenberg, 1997; Greeno & Moore, 1993; Kosslyn, 1999; Lakoff & Johnson, 1999; Wexler et al., 1998; and in new media studies, e.g., Burnett, 2004; Hansen, 2004) as something relative to an internal (neurological or psychological) process of an individual. At the same time it is well known that social and cultural factors are constitutive for every visual experience (e.g., Eco, 1997; Deleuze, 1986). The centrality of social and cultural elements in scientific practice is clearly visible in the analysis of the fMRI article (Chapter 2). The analysis illustrates how scientists progressively negotiate the appearance of brain images with the standards of the neuroscience community.

However, the analysis limited to an examination of written documents leaves the full richness of the embodied practice largely unexamined (other than pointing out the centrality of visual representations for brain mapping practice). The question is then: how can one reconcile embodiment with the involvement of social and cultural factors in visual experience?

To deal with the problem the present work is attentive to micro-aspects of the everyday life-world of MRI brain mapping. Such a focus indicates not only how the body of an individual is drawn into processes of interpretation, but also how interpretation regards the human body implicated in social action (this argument is developed through Chapters 3-5). Gesture, for instance, allows not only its producers, but also their interlocutors to engage with embodied gestural performances. In other words, an understanding of vision as an embodied process, which involves coordination of semiotic modalities with the material and social world of action, indicates that visual processes are embodied processes situated in the interaction between social actors (see also Goodwin, 2003c).

Understanding digital worlds of images as enacted, publicly shared and mutually co-dependent semiotic fields

The abandonment of a language-centered analysis (or any analysis focused on just one semiotic modality, such as gesture – e.g., McNeill, 1992) permits a novel way of theorizing about communication and interaction. It allows the interdependencies between action and meaning-making processes to emerge.

As previously remarked, MRI practitioners point at, reach toward, and touch computer screens. Through these actions, they not only index information on the digital screens, but also “pull” the information from the screens into the 3D world of public action. Once information is translated into the world of action, it can be further manipulated, performed and acted upon. While enacting certain aspects of an image through gesture, practitioners create new renderings of the sheared image (Chapter 5). By enacting and integrating such *semiotic fields* the practitioners can instantiate novel ways of looking at the problem at hand by, for example, performing “sheared brain” and the process of its shearing.

Embodied semiotic fields are short-term scenarios constructed on-line. Like Fauconnier and Turner’s (ibid.) mental spaces, the semiotic fields can, for example, be quickly rearranged, and can instantiate imaginary elements. Whereas it is unlikely to see an actual brain being sheared (since the “sheared” aspect of the digital image was not actually caused by any physical action upon the brain), or to observe such a process by looking at a single, static digital brain representation, to observe such an action while looking at the gesture interacting with other semiotic resources while performing shearing is a practical way of dealing with experimental data.

Unlike internal mental spaces, semiotic fields are material representations that exist in the external world (i.e., *material anchors*, Hutchins, in press), or dynamic enactments that develop through action. Such enactments involve the body and provide collective access to information. By observing gestural performance, gaze direction, and body positioning and orientation, MRI practitioners can negotiate and coordinate their looking of a brain representation. They can frame (e.g., through body positioning) and index (e.g., through gaze and body orientation) elements of the world that need to be focused on, and expose particular elements of the scene for collective viewing. The timing of the instantiations of semiotic fields is also important: it enhances the possibility of their collective production (e.g., by slowing down the gestural performance, a practitioner can allow her interlocutor to co-participate in the process). Furthermore, while engaging with embodied semiotic fields, practitioners can manipulate rather complex structures, such as the detailed and colorful brain maps. They can also deal with multiple formats at once (e.g., similar information can be displayed in the form of an fMRI brain image, as well as in the form of a hand-drawn chart representing human cortical organization, cf. Chapter 4). This multiplicity of formats enhances possibilities of exploration (Goldin-Medow, unpublished manuscript), sharing and embodied involvement with abstract information.

Action as cognition

In addition to recognizing the centrality of embodied semiotic fields for meaning-making processes, observations of fMRI practice reveal how such fields interact and are linked together. It has been claimed that the capacity to integrate disparate domains into composite conceptual structures is characteristic of human thinking (e.g., Fauconnier & Turner, ibid.; Koestler, 1964; Mithen, 1998). Here such processes are being executed in the external world of practical action. Engagements with digital screens and other social actors do not only *reflect* cognitive processes (Chapter 4) (e.g., Goldin-Medow et al., 2001; Garber et al., 2002). Instead, such engagements are, themselves, cognitive processes (for a comparable view see also Goldin-Medow, unpublished abstract). Once we start considering scientists' performances in terms of the enactments of semiotic fields and the processes of their coordination, we notice how social, material, and embodied actions produce an understanding of visually expressed information. Scientists act in social settings to generate meanings of visual displays. In other words, they use their hands and bodies to make sense of the material world through social actions (Latour, 1986).

The interaction among semiotic fields generates unexpected phenomena. The apparent incompatibility between fields, as well as between elements and processes inside such fields, often generate *fantastic* events. Observations of multimodal interaction allow us to see processes of blending and integration between semiotic fields as they unfold in the external world of action. For example, the enactment of “brain shearing” (Chapter 5) allows scientists to conceive of multiple semiotic fields - the physical brain, its visual representation on the computer screen, and abstract data that such a representation stands for – to be manipulated as if they were a single entity. The integration of multiple fields produces a multifaceted form that allows the co-existence of an *impossible* combination of characteristics. The stucture has, at the same time, some of the properties of physical object (i.e., it can be directly manipulated), while it is conceived as being a digital image and experimental MRI data. These processes of integration between semiotic fields are observed by scientists while they employ them for sense-making and practical problem solving. Even though scientists obviously know that the account of brain image artifact in terms of an act of physical shearing is, in fact, *incorrect*, the extremely complex and fantastic process constructed through the embodied and social action is a more concrete (see also Lakoff, 1987; Fauconnier & Turner, 2002) and practical way of dealing with experimental data. It immerses semiotic activities in the flow of laboratory work (Chapter 5).

DCog as computation within a functional system vs. dynamic and embodied semiotic performances

To propose an existence of external cognitive processes does not imply a negation of their internal counterparts. A complete account of cognition in scientific research (or any other setting) would include, in addition to the analysis of processes that occur in the interactions of individuals with one another and with the material environment of action, a careful description of internal cognitive processes. 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 observed in my research setting.

While research in cognitive science has been mainly focused on such internal processes, the knowledge of social coordination has been domain of social science, where minute details of embodied practice and its cognitive aspects have been largely overlooked. Since the two types of processes – the internal brain processes, and the external social, cultural, and embodied processes - are strictly co-dependent, hypotheses concerning what people do with their minds must be grounded in a specification of the embodied practices that individuals perform in socio-cultural and material settings. An understanding of the way in which individuals engage with their environments and other actors can only be achieved by carefully examining the organization of action in that setting.

To study a social system as the unit of cognitive analysis, Ed Hutchins (1995) proposed to describe the activity of the system in the way cognition has been traditionally described. The activity is seen as a computation realized through the creation, transformation, and propagation of representational states across media within a functional system (ibid.: 49; 372-3). But, as pointed out by Hutchins, a distributed cognition perspective doesn't need to commit to a narrow view of computation (ibid.: 118).

In this respect, the present work expands on and moves beyond the characterization of activity strictly in terms of computational processes within a functional system. In describing the internal workings of the human mind contemporary theories in cognitive science have emphasized the importance of action, intention, emotion, and bodily grounded experience (e.g., Freeman & Núñez, 1999; Ballard et al., 1997; Fuster, 1997; Carpenter et al., 2000; Jeannerod et al., 1995). By adopting these principles, as well as the principles of *conceptual integration theory*, and applying them to the workings of socio-cultural and technological systems, the present study calls attention to highly dynamical processes of interpretation, coordination, co-construction and mutual elaboration between semiotic fields.

Inasmuch it concerns semiotic processes in the context of science, the study is closely related to Latour and Woolgar's (ibid.) study of the neuroscience laboratory. However, while Latour and Woolgar's base their semiotics in *structuralist* tradition and the formal approach to narrative of Algirdas Julien Greimas (Greimas & Courtes, 1983), the present work emphasizes embodied and dynamic processes of *semiosis*, having its point of departure the *interpretative semiotics* of Charles Sanders Peirce (1934-48). Peircian semiotics, with its emphasis on the in-built dynamism of signs - where each sign necessarily generates further signs and functions as a manifestation of a general phenomenon (for a more detailed exposition see, for example, Alač, 2002) - allows for a direct engagement with the “messy” and “disordered” world of embodied routines. Rather than “reconfiguring local expressions into various decontextualized representations” (Lynch, 1993: 310), general postulates of interpretative semiotics (as well as borrowings from *conceptual integration theory*) allow for “messy” and “disordered” actions of laboratory work to be *recovered*. Thus, semiotic performances are not *delocalized* to the *analytically transparent* networks of signs (ibid.: 292), but feature in the phenomenal field of laboratory praxis. Their identification is itself a situated activity that relies on specific details of practical action (Suchman, 2000:17).

DISSERTATION ORGANIZATION

In what follows, I will briefly outline the organization of the entire document. The text (Chapters 2-5) traces the role of brain images by describing MRI practices that range from publishing a scientific report to teaching the novice how to analyze experimental data. This organization does not reflect a chronological order in which images “migrate” from the scanning room, to the laboratory, and into the public domain of published results. Instead, the exposition moves from the less controversial claim that visual experiences reflect cultural and social practices of standardization, to the descriptions of the embodied character of such experiences and their consequences for problem solving and knowledge acquisition as material, embodied, real-time, practical accomplishments.

CHAPTER 2: NEGOTIATING IMAGES – CULTURAL AND SOCIAL ASPECTS OF BRAIN IMAGES; STANDARDS

In this chapter I examine scientific documents in order to describe a process of submitting, revising, and publishing an experimental research article in cognitive neuroscience.

The review process, as part of scientific practice, is of particular interest, since it puts the research team in direct dialog with a larger scientific community concerned with fMRI evidence. By bringing this often ‘black-boxed’ dimension of the manuscript’s production into the picture, I illustrate the role played by the visual brain representations in the practice of scientific fact production. This allows me to show how the knowledge is constructed through seeing, but also to point out how the showing of the fact through images bears traits of the scientific culture of which it is part.

In cognitive science it is by now an unquestioned fact that visuospatial representations are particularly apt for problem solving (e.g., Larkin & Simon, 1987; Glasgow, Narayanan, Chandrasekaran, 1995; Anderson, Meyer, Olivier, 2002; Kirsh, 1995). Moreover, proponents of cognitive semantics and embodiment (e.g., Lakoff, 1987; Fauconnier & Turner, 2003) have indicated that understanding of abstract entities and processes is often based upon biological capacities and experiences of functioning in a physical and social environment. Since humans use vision as our primary sense, it is cognitively more apt to analyze abstract entities (e.g., experimental data expressed in numerical terms) as visual and spatial (e.g., brain images).

Yet, to explain the fact that visualization deeply shapes scientific activity in terms of the processes relative to individual organisms does not entirely do justice to the reality that emerges from the revision processes described in this chapter. The analysis emphasizes the centrality of the brain images in the production of scientific fact. Nevertheless, it also illustrates the inevitable embeddedness of scientific text in a particular social, cultural, and historical context, as well as its dependence on the audience. These elements are more easily overlooked when one is immersed in the ‘clean’ world of experimental procedures. However, they become strikingly real in scientists’ everyday practices, such as the revision process.

One of the symptomatic points that came up through the revision process was the reviewers’ dissatisfaction with the appearance of ‘regions of activity beyond the two that the manuscript focused on’. The reviewers’ remarks implied inclinations and expectations to see specific locations of activation, as is customary in fMRI literature. As this negotiation of the ‘number of cortical activation spots’ shows, the link between the image and its referent is always mediated through the filters of social and cultural norms and expectations. ‘Seeing’ of reality becomes ‘clearer’, ‘less confusing’, and ‘easier to verify’ when it follows the parameters of the scientific culture of which it is part.

This chapter is based on Alač, "Negotiating pictures of numbers", *Journal of Social Epistemology, Journal of Social Epistemology*, 18:2, 2004, 199-214.

CHAPTER 3: SENSE MAKING THROUGH COORDINATION OF MULTIPLE SEMIOTIC MODALITIES IN MRI PRACTICE

While the first chapter focuses on written documents, this chapter takes in consideration particulars of embodied practice. I show how coordination of multimodal semiotic modalities reveals the way in which scientists make sense of central aspects of their practice such as image artifacts.

In order to produce brain images that can be successfully analyzed, cognitive neuroscientists design and carry out fMRI experiments where movement of the subject during the scanning session as well as the high decibel acoustic noise produced by the scanner need to be reduced to a minimum. Movements and noise are very harmful for the image production; they are nothing more than obstacles for research.

While observing instances of fMRI practice, I, however, focus on these *obstacles*, rather than disregard them. I look at how such phenomena are reenacted as part of discursive social action and what their role is in knowledge production. I show how these unwanted aspects of the scientific procedure are reused by researchers and made functional.

In order to trace ways in which such elements *leak back* into the process of scientific sense making, I focus on traditionally less explored communication modalities, such as non-linguistic sounds, body movements, gestures, etc.. I describe how such modalities reintroduce the unwanted elements of the scientific practice into the processes of its understandings. The understanding is achieved by coordination of various semiotic modalities with brain images. In other words, the analytic focus on the use of and coordination between different semiotic modalities allows us to see how scientific practice develops as an embodied, social activity.

The *reuse* of the motion and noise for understanding appears in two instances of apprenticeship. In the first example, the novice acquires the categories of the profession (Goodwin, 1994) by listening to the expert's reenactment of the scanner's noise. The confounding factor becomes an important element in the embodied procedure of sense-making. The expert's vocalizing works as a *coding schema* which organizes the messy world into categories relevant for the work of the profession (ibid.). The construction of knowledge and rationality is neither linguistic nor graphical in form; rather it is a nonlinguistic, acoustic event (for a discussion of the use of sound in science see Cartwright, in preparation). The second example regards the *reuse* of motion. Participants reason about the nonalignment between brain images through perceiving and re-performing the hypothetical movement of the experimental subject. They construct the subject's movement through the coordination of language and brain images, with gesture and bodily movements. In such a way, that which can only be imagined[3] is made visible in the shared space of action.

My analysis of the embodied apprenticeship practice exposes the way in which the concept of motion artifact gets constructed. The artifact becomes visible neither on the computer screen nor in the movement that the participants have to imagine. Instead, participants create a hybrid construct where the properties of the subject movement are integrated with properties of the brain images. The artifact is not a thing or a well-defined location (Lynch, 1985). Rather, it reveals itself in the linkage generated in the public space of practice among various semiotic modalities.

---

[3] The movement of the subject could not have been seen or experienced by the expert and the novice while the subject was laying in the scanner; the movements that the subject produces during the scanning section are usually too small to be seen on the computer monitor where fMRI technicians and researchers supervise what goes on in the scanner.

This chapter is based on Alač, "From Trash to Treasure: Learning about the Brain through Multimodality", *Semiotica*, 156-1/4, 2005, 177-202.

CHAPTER 4: ACTION AS COGNITION

In this chapter I further elaborate on the idea of cognition as a process largely enacted in the world of social and embodied action.

We like to consider us – humans – as thinking creatures par excellence. It has been claimed that a definite trait of our thinking skill is the ability to combine different conceptual domains (e.g., Fauconnier & Turner, ibid.; Koestler, 1964; Mithen, 1998).

By observing how MRI practitioners engage with digital screens, I notice that acts of thinking and coordination of diverse conceptual domains do not only take place “in the head” of an isolated thinker. Such acts can be accomplished through embodied and social actions. Practitioners individuate meaningful entities within fMRI representations by coordinating brain images with other cultural and embodied representations distributed across the environment of practice. This type of thinking is denominated “action as cognition”.

In this chapter I describe an example of such a process where an experienced practitioner explains to a novice the organization of a visual brain area (represented via an fMRI image). To do this, she identifies a mapping between the locations in visual field (sketched on a piece of paper) and the locations in visual brain area. In other words, the practitioners need to understand the organization of the fMRI experimental data present on the computer screen by putting it in correspondence with a sketched diagram of the visual field. To do so, the diagram of the visual field provides an anchor for an imagined visual field object. Such an object can be subjected to an embodied action of squeezing. By engaging their hands with the visual representation, scientists create the semiotic object (i.e., the visual field), and enact an embodied act of squeezing on such an object. This act is employed to translate one semiotic entity in the other: the visual field is transformed into its representation on the human cortex. The two objects employed in the enactment do not independently exist in the environment. Instead, scientists create such entities through drawing and gesturing. What is even more surprising, the scientists map one object or semiotic domain onto the other by enacting such process in the environment of practice. The mapping and transformation of one entity into the other is fictional in the sense that it represents the change as a physical process - squeezing, when, in fact, the transformation is produced in the visual system by the structure of neural circuitry. At the same time, the process is constructed in front of the researchers' eyes. This type of action constitutes thinking for the researchers. The researchers both produce the meaning of scientific inscriptions and understand the meaning that is produced by performing these actions. They employ their bodies to find out how the two conceptual fields are associated, and how the brain image is organized.

In addition to talk, gesture is an especially important feature of action as cognition. Gesture not only conveys or encodes meaning; it often ties together different semiotic fields. Gesture can link distinct fields into larger schematic units, highlighting potential, but otherwise unmarked, dependences among the fields. For example, the "squeezing the visual field" gesture builds a link between the visual field representation and the digital representation of the brain cortex by creating a dynamic process. Importantly, these relationships and linkages are performed in the public space of action. The problem of linking disparate domains into composite structures is one of the central questions in cognitive science (for discussion see Fauconnier & Turner, ibid.). Here such a cognitive process is being executed in the external, shared world or practice.

This chapter is based on Alač & Hutchins: "I See what You are Saying: Action as Cognition in fMRI Brain Mapping Practice", *Journal of Cognition and Culture*, 4:3, 2004, 629-661.

CHAPTER 5: SEEING DIGITAL BRAIN IMAGES AS HYBRID OBJECTS TO WORK WITH

Whereas the two previous chapters focus on apprenticeship activities, this chapter centers on an interaction between two experienced MRI researchers (the PI and the PD). The researchers work in two different laboratories that study the human visual system. The excerpt from the interaction analyzed here portrays a working session where the researchers participate in an analysis of the MRI data that were previously analyzed by the PD. The goal of the re-analysis is to compare the software packages designed and used in the two laboratories. During the interaction the PI notices a “shearing” in the functional image that hasn’t been noticed when the data were analyzed in the other laboratory. In brain mappers’ jargon, “shearing correction” indicates a group of computational processes performed on images in order to correct a particular type of distortion. Similarly, the distorted image is described as being “sheared”. As he tries to correct the image distortion, the PI first demonstrates to the PD how a “sheared” image is identified, and then how its correction is achieved through the use of the software adopted in his laboratory.

While focusing on the activity of shearing correction, the chapter resumes and further explores some of the main topics of the dissertation.

First, it looks at how the relevant feature of the image is made collectively visible through its instantiation in the public world of action. The social space, rather than the image space, is the primal site of the production of visibility. The excerpt shows how gesture plays a crucial role in such a production. The gestural performance transforms the physical space shared by the practitioners into a *knowledge space*.

What is particularly important about the gestural performance is its coordination with other semiotic structures that participate in the meaning-making processes. Rather than simply functioning as translations of the *world of sight* into the *world of touch* (Streeck, in press), the coordination allows for a *permeability* between the digital and the physical worlds.

The chapter illustrates how embodied semiotic means participate in the coupling between the computer screen and the action instantiated in the physical space.

The integration of the physical and virtual space generates phenomenal entities which permit the co-existence of a *rationally impossible* or *counterfactual* combination of characteristics. One such entity may have some of the properties of the physical object (i.e., it can be directly manipulated), while being conceived as a digital image. The analysis of micro-aspects of the interaction and work reveals that the production of such entities is accomplished through the coordination of the digital world with multimodal semiotic resrouces in the environment of MRI practice. The PI reaches toward the computer screen and gestures as if holding a round three dimensional object and moving it towards left/right. Similarly, the PI performs the action of "squashing" over the image, as if squeezing a three-dimensional object. The PI's gestures are accompanied by linguistic expressions (i.e., "If you rotate 30 degrees this way"; or "and then squash") that help disambiguate the gestural form, and characterize it in terms of physical manipulations. On the other hand, the gesture's shape and size is determined by the image: the hand almost wraps around the brain image and follows its contours with precision. While the gesture performs the action of 3-D physical manipulation, its form is shaped by the appearance of the digital data. The gestural capacity to coordinate multiple domains is enhanced by its precise spatial location.

The existence of such a coordination that generates hybrid and fantastic entities leads one to wonder if there is any relationship between such entities and the concrete patters of laboratory work. By focusing on the minute instances of the discursive actions, the chapter discusses how such culturally and socially situated, embodied acts of meaning-making concern practical ways in which professional neuroscientists deal with their experimental data.

CHAPTER 6: CONLCUSION - DISTRIBUTED COGNITION, EMBODIMENT AND MULTIMODAL INTERACTION IN MRI PRACTICE

The chapter discusses the relationship between the present work and embodied and distributed cognition frameworks.

WORK CITED

ALAČ, M.
2002 *Interpretative Semiotics and Creative Mind*, Unpublished doctoral dissertation, University of Bologna, Italy.

ANDERSON, M., MEYER, B., OLIVIER, P. (eds.)
2002 *Diagrammatic Representation and Reasoning*, London: Springer-Verlag.

BAKHTIN, M.
1988 *Speech Genres and Other Late Essa*y, Austin: University of Texas Press.

BALLARD, D.H., HAYHOE, M.M., POOK, P.K., RAO, R.P.N.
1997 "Deictic codes for the embodiment of cognition", *Behavioral and Brain Sciences* 20: 723-767.

BARSALOU, L.W.
1999 "Perceptual symbol systems", *Behavioral and Brain Sciences* 22: 577-660.

BARTHES, R.
1988 (1985) *The Semiotic Challenge*, Berkeley and Los Angeles: University of California Press.

BEAULIEU, A.
2002 “Images are not the (only) truth: Brain mapping, visual knowledge, and iconoclasm”, *Science; Technology, & Human Values* 27, 53-86.

2004 “From brain bank to database: the informational turn in the study of the brain”, *Studies in History and Philosophy of Biological and Biomedical Sciences* 35: 367-390.

BURNETT, R.
2004 *How Images Think*, Cambridge, MA: MIT Press.

CARPENTER, P., JUST, M.A., REICHLE, E.D.
2000 "Working memory and executive function: Evidence from Neuroimiaging", *Current Opinion in Neurobiology* 10: 195-199.

CARTWRIGHT, L.
In preparation “An Industry Melodrama: Reading US Motion Picture Engineers’ Writings on Film Sound Technology, 1930-1950”.

CARTWRIGHT, L. & STURKEN, M.
2001 *Practices of Looking: An Introduction to Visual Culture*, Oxford: Oxford UP.

CLARK, A.J.
1987 Being there: Why implementation matters to cognitive science. *Artificial Intelligence Review* 1: 231—244.

1997 *Being There: Putting Brain, Body, and World Together Again*. Cambridge, MA: MIT Press/London: Bradford Books.

2001 *Mindware: An Introduction to the Philosophy of Cognitive Science*, Oxford: Oxford UP.

CLARK, A. & THORNTON, C.
1997 "Trading Spaces: Computation, Representation, and the Limits of Uninformed Learning", *Behavioral and Brain Sciences* 20, 57-90.

COLE, M.
1998 *Cultural Psychology: A Once and Future Discipline*, Cambridge, MA: Harvard UP.

COULSON, S.
2000 *Semantic Leaps: Frame-shifting and Conceptual Blending in Meaning Construction*, Cambridge: Cambridge UP.

COULSON, S. & OAKLEY, T.
2000 "Blending Basics", *Cognitive Linguistics* 11, 3-4: 175-196.

DELEUZE, G.
1986 (1983) *Cinema One: The Movement-Image*, Minneapolis: University of Minnesota Press.

DUMIT, J.
2004 *Picturing Personhood: Brain Scans and Biomedical Identity*, Princeton, NJ: Princeton UP.

ECO, U.
1976 (1975) *A Theory of Semiotics*, Bloomington: University of California Press.
1997 *Kant and the Platypus: Essays on Language and Cognition*, New York: Harcourt and London: Secker.

ENGESTROM, Y. & MIDDLETON, D. (eds.)
1996 *Cognition and Communication at Work*, New York: Cambridge UP.

FAUCONNIER, G.
1994 *Mental Spaces*, Cambridge: Cambridge UP.

FAUCONNIER, G. & TURNER, M.
1998 "Conceptual Integration Networks", *Cognitive Science*, 22, 133-187.

2002 *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*, Basic Books.

FRIEDBERG, A.
2003 "The virtual window", In *Rethinking Media Change: The Aesthetics of Transition*, D. Thorburn & H. Jenkins (eds.), Cambridge, MA: MIT Press.

FUSTER J.M.
1997 "Network memory", *Trends in Neuroscience* 20: 451-459.

GALISON, P.
1997 *Image and Logic: A Material Culture of Microphysics*, Chicago: Chicago UP.

GARFINKEL, H.
1967 *Studies in Ethnomethodology*, Englewood Cliffs, NJ: Prentice-Hall.

GEERTZ, C.
1973 *The Interpretation of Cultures*, New York: Basic Books.

GIERE, R.
In press “The Problem of Agency in Scientific Distributed Cognitive Systems”, *Journal of Cognition and Culture*.

GLASGOW, J., NARAYANAN, N. HARI, CHANDRASEKARAN, B. (eds.).
1995 *Diagrammatic Reasoning: Cognitive and Computational Perspectives*, Cambridge, MA: MIT Press.

GLENBERG, A.
1979 "What memory is for", *Brain and Behavioral Sciences* 20(1): 1-55.

GREIMAS, A.J. & A. COURTES
1983 (1979) *Semiotics and Language: Analytical Dictionary*, Bloomington: Indiana UP.

GOFFMAN, E.
1967 *Interaction Ritual: Essays on Face-to-Face Behavior*, New York: Doubleday Anchor.

GOLDIN-MEDOW, S.
Unpublished abstract "Hearing gesture: how our hands help us think", UCSD *Cognitive Science 200 Seminar on Gesture*, *Spring 2004*.

GOODWIN, C.
2003a "Pointing as Situated Practice", IN S. Kita (ed.), *Pointing: Where Language, Culture and Cognition Meet*, Mahwah, NJ: Lawrence Erlbaum.

2003b "Conversational Frameworks for the Accomplishment of Meaning in Aphasia", IN C. Goodwin (ed.), *Conversation and Brain Damage*, Oxford: Oxford UP.

2003c "The Semiotic Body in its Environment", IN J. Coupland & R. Gwyn *Discourses of the Body*, New York: Palgrave/Macmillan.

2000a "Action and Embodiment Within Situated Human Interaction", *Journal of Pragmatics* 32: 1489-522.

2000b "Gesture, Aphasia and Interaction", IN D. McNeill (ed.), *Language and Gesture: Window into Thought and Action*, Cambridge: Cambridge UP.

2000c "Practices of Seeing, Visual Analysis: An Ethnomethodological Approach", IN T. van Leeuwen & C. Jewitt (eds.), *Handbook of Visual Analysis*, London: Sage.

1997 "The Blackness of Black: Color Categories as Situated Practice", IN L. Resnick, R. Säljö, C. Pontecorvo and B. Burge (eds.), *Discourse, Tools and Reasoning: Essays on Situated Cognition*, New York: Springer-Verlag.

1996 "Seeing in Depth", *Social Studies of Science* 25: 237-74.

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

GOODWIN, C., GOODWIN, M.H., and M. YAEGER-DROR
2002 "Multi-modality in Girls' Game Disputes", *Journal of Pragmatics* 24 (10-11): 1621-49.

GREENO, J. & MOORE, K.
1993 "Situativity and symbols: Response to Vera & Simon", *Cognitive Science* 17 (1): 49-59.

HANSEN, M.B.N.
2004 *New Philosophy for New Media*, Cambridge, MA: MIT Press.

HAYLES, K.N.
1999 *How we became Posthuman: Virtual Bodies in Cybernetics*, Literature and Informatics, Chicago: University of Chicago Press.

HEATH, C. & J. HINDMARSH
2002 "Analysing Interaction: Video, ethnography and situated conduct" In T. May (ed.) *Qualitative Research in Action*, London: Sage: 99-121.

HOLLAN, J., HUTCHINS, E., KIRSH, D.
2000 "Distributed Cognition: Toward a new foundation for Human-Computer interaction research", ACM Transactions on Computer-Human Interaction, 7.

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

1999 "Cognitive Artifacts", MIT Encyclopedia of Cognitive Science.

2005 "Material anchors for conceptual blends", *Journal of Pragmatics* 37/10: 1555-1577.

HUTCHINS, E. & PALEN, L.
1993 "Constructiong Meaning From Space, Gesture, and Talk", In *Discourse Tools and Reasoning*, (L.B. Resnick et.al. eds.), Springer Publishing.

JACK, A.I. & ROEPSTORFF, A.
2002 "Introspection and cognitive brain mapping: from stimulus-response to script-report", *Trends in Cognitive Science*, 6, 333-339.

JEANNEROD, M., ARBIB, M.A., RIZZOLATTI, G. SAKATA, H.
1995 "Grasping Objects: The Cortical Mechanisms of Visuomotor Transformation", *Trends in Neurosciences* 18 (5): 314-320.

KARMILOFF-SMITH, A.
1992 *Beyond Modularity*, Cambridge, MA: MIT Press.

KENDON, A.
1990 *Conducting Interaction: Patterns of Behavior in Focused Encounters*, Cambridge: Cambridge UP.

KIRSH, D.
1995 "The Intelligent Use of Space", *Artificial Intelligence* 73, 31-68.

KNORR-CETINA, K.
1981 *The Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science*, Oxford: Pergamon Press.

1999 *Epistemic Cultures: How the Sciences Make Knowledge*, Cambridge, MA: Harvard UP.

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

KOSSLYN, S.M. et al.
1999 "The Role of Area 17 in Visual Imagery: Convergent Evidence from fMRI and rTMS", *Science* 284(5411): 167.

KUHN, T.
1962 *The Structure of Scientific Revolution*, Chicago: University of Chicago Press.

LAKOFF, G.
1987 *Woman, Fire, and Dangerous Things*, Chicago: Chicago UP.

LAKOFF, G. & JOHNSON, M.
1999 *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*, Basic Books.

LANGACKER, R.W.
1991 *Concept, Image, and Symbol : The Cognitive Basis of Grammar*. (Cognitive Linguistics Research, No 1) Mouton De Gruyter.

LARKIN, J. & SIMON, H.
1987 “Why a diagram is (sometimes) worth ten thousand words”, *Cognitive Science* 11:65-99.

LAVE, J.
1988 *Cognition in Practice: Mind, mathematics and culture in everyday life*, Cambridge: Cambridge UP.

LATOUR, B.
1986 ‘Visualization and Cognition: Thinking with Eyes and Hands’, *Knowledge and Society* 6: 1-40.

1987 *Science in Action*, Cambridge, MA: Harvard UP.

LATOUR, B. & WOOLGAR, S.
1979 *Laboratory Life. The Social Construction of Scientific Facts*, London and Beverly Hills: Sage.

LENOIR, T.
2004 “Foreword” In M.B.N. Hansen.

LOTMAN, Y.
1990 Universe of the Mind: A semiotic theory of culture, Bloomington: Indiana University Press.

LYNCH, M.
1985 Art and Artifact in Laboratory Science: A Study of shop work and shop talk in a research laboratory, London: Routledge & Kegan Paul.
1993 Scientific Practice and Ordinary Action: Ethnomethodology and social studies of science, Cambridge: Cambridge UP.

MAC EACHREN, A.
1995 *How Maps Work: Representation, Visualization, and Design*, New York: The Guilford Press.

MANOVICH, L.
2001 *The Language of New Media*, Cambridge, MA: MIT Press.

McNEILL, D.
1995 *Hand and Mind: What Gesture reveal about Thought*. Chicago: University of Chicago Press.

McNEILL, D. (ed.)
2000 *Language and Gesture*. Cambridge: Cambridge UP.

MIDDLETON, D. & EDWARDS, D.
1990 *Collective Remembering*, London: Sage.

MINDELL, D.
2002 *Between Human and Machine: Feedback, Control, and Computing Before Cybernetics*, Baltimore: John Hopkins UP.

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

MIRZOEFF, N.
1999 *An Introduction to Visual Culture*, London and New York: Routledge.

NEISSER, U.
1982 *Memory Observed: Remembering in natural contexts*, Oxford: Freeman.

NORMAN, D.A.
1988 *The Psychology of Everyday Things*, New York: Basic Books.

1991 “Cognitive artifact” IN J. M. Carroll (ed.), *Designing interaction*, Cambridge, MA: Cambridge UP.

NUNEZ, R. & FREEMAN, W. (eds.)
1999 *Reclaiming Cognition: The Primacy of Action, Intention and Emotion, Journal of Consciousness Studies*, Vol. 6, November/Dicember.

PEIRCE, C. S.
1934 -1948 *Collected Papers*, Cambridge, MA: Harvard UP.

PICKERING, A.
1995 *The Mangle of Practice: Time, agency, and science*, Chicago: University of Chicago Press.

PULVERMULLER, F.
1999 "Words in the Brain's Language", *Behavioral and Brain Sciences* 22: 253-336.

ROEPSTORFF, A.
2001 “Brains in scanners: an Umwelt of cognitive neuroscience”, *Semiotica*, 134, 747-765.

2003 “A double dissociation in twentieth century psychology?”, *Journal of Consciousness Studies*, 10: 62-67.

2004 “Mapping brain mappers: An Ethnographic Coda”, IN R.S. Frackowiak et al. (eds.), Human brain function (2nd ed.), Amsterdam: Elsevier.

ROEPSTORFF, A. & FRITH, C.
2004 “What’s at the top in the top-down control of action? Script-sharing and ‘top-top’ control of action in cognitive experiments”, *Psychological Research* 68: 189-198.

ROGOFF, B. & LAVE, J.
1984 *Everyday Cognition: Its development in social context*, Cambridge, MA: Harvard UP.

SERENO, M., PITZALIS, S., MARTINEZ, A.
2001 "Mapping of Contralateral Space in Retinotopic Coordinates by a Parietal Cortical Area in Humans", *Science* 294: 1350-1354.

SERENO, M.I., DALE, A.D., REPPAS, J.B., KWONG, T.J., BELLIVEAU, J.W., BRANDY, T.J., ROSEN, B.R., TOOTELL, R.B.H.
1995 "Borders of multiple visual areas in humans revealed by functional MRI", *Science*, 268: 889-893.

SUCHMAN, L.
1987 *Plans and Situated Actions*. Cambridge: Cambridge UP.

STREECK, J.
(In Press) “Pragmatic Aspects of Gestures”, *The Encyclopedia of Language and Linguistics*, Keith Brown (ed.), Oxford: Elsevier.

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

TOOTELL, R.B.H., HADJIKHANI, N.K., VANDUFFEL, A.K.L., MENDOLA, J.D., SERENO, M.I., DALE, A.D.
1998 "Functional analysis of primary visual cortex (V1) in humans", *Proceedings of National Academy of Science, USA*, Vol.95: 811-871, February.

TUFTE, E.R.
1983 *The Visual Display of Quantitative Information*, Cheshire, Conn.: Graphics Press.

WANDELL, B.A.
1999 "Computational neuroimaging of human visual cortex", *Annual Review of Neuroscience* 22: 145-73.

WERTSCH, J.
1998 *Minds as Action*, New York: Oxford University Press.

WEXLER, M., KOSSLYN, S.M., and BERTHOZ, A.
1998 "Motor processes in mental rotation", *Cognition* 68: 77-94.

CHAPTER 2

NEGOTIATING PICTURES OF NUMBERS

ABSTRACT

This paper reports on objectivity and knowledge production in the process of submitting, revising, and publishing an experimental research article in cognitive neuroscience. The review process, as part of scientific practice, is of particular interest, since it puts the research team in direct dialog with a larger scientific community concerned with fMRI evidence. By bringing this often ‘black-boxed’ dimension of the manuscript’s production into the picture, I illustrate the role that the visual brain representations played in the practice of scientific fact production. This allows me to point out how the knowledge is constructed through seeing, and how the showing of the fact through images bears traits of the scientific culture of which it is part.

1. INTRODUCTION

An experimental research article is the end product of laboratory work. At the same time, however, it is also the result of the writing and revision process. It distils the laboratory work, while it silently represents practices of negotiation with the larger scientific community. These negotiations are clearly manifested through the publishing process.

Since the publishing process reveals the interface between the local research and the larger scientific community, its description and analysis should be a significant component of the study of objectivity. Yet, even though important work has been done in this domain[4], the publishing process has received less attention than it deserves. This lack is mainly due to its impenetrability.

While conducting my ethnographic research in a neuroscience laboratory, I was given access to documentation that testifies to production of scientific writing. The writing in question is an experimental research article that was recently published in Science (Sereno et al., 2001). In my analysis I look at its process of submission, revision, and publishing. At the same time I focus on the status of the brain images and their function in the article. I analyze the way in which objectivity and knowledge, relative to brain images, are culturally negotiated.

As a member of the cognitive science community, I enter the laboratory as a ‘native’ of the community. At the same time I conduct an observational analysis of the ways in which scientific evidence and knowledge are produced using visual representations (e.g., Alač and Hutchins, 2004). In this sense, I occupy the dual role of cognitive scientist and social science investigator focusing on the communication and knowledge in my research setting[5].

The interest in the role of visual and visuo-spatial representations in scientific knowledge production is not new. Research in cognitive science, for instance, has been for quite some time concerned with the role of visual representation in cognition (e.g., Glasgow, Narayanan, Chandrasekaran, 1995; Anderson, Meyer, Olivier, 2002). Yet, the interdependencies between such representations and cultural and social processes have been largely left unnoticed. On the other hand, researchers committed to investigate the inseparability of visual representations in science from social and cultural practices (e.g., Lynch & Wolgar, 1990) have not been particularly keen on adopting cognitive science

---

[4] Charles Bazerman (1988) studied the genre of experimental research article as a cultural form, and traced its emergence, evolution and stabilization. Karin Knorr-Cetina (1981) looked at how a scientific article, while reporting on the laboratory research, filters what happened in the laboratory. Knorr-Cetina also illustrated how scientific article employs literary strategies.

[5] For a discussion on problems inherent in studying one’s own environment, and possible advantages of such endeavour see Bourdieu, 1984.

terminology. I believe that borrowing and contamination across the two research domains could be productive. In the discussion of the publishing process I employ the vocabulary of cognitive science to point out the cultural and social character of the phenomena that such vocabulary refers to.

2. ETHNOGRAPHY OF FMRI IMAGES

Cognitive neuroscientists use Functional Magnetic Resonance Imaging (fMRI) technology to produce images of active processes within the brain. These images are generated by visualizing the local changes in magnetic field properties occurring in the brain. Such changes are a consequence of alterations in blood oxygenation. If the image is obtained while an experimental subject is engaged in a particular cognitive task, the image can indicate which parts of the brain are most active in that task.

The present analysis is a part of a broader ethnographic study on the production of visual fMRI evidence. The study involves observations of scientific practices conducted over a period of 9 months in three laboratories. Data collection methods include direct observation, video, semi-structured interviews, and analysis of documents such as journal articles, laboratory manuals and scientific correspondence.

The first author of the research article under examination is the principal investigator (PI) of one of the fMRI laboratories where I conduct my ethnographic research. The laboratory researches human visual perception through fMRI technique. The primary interest is to localize particular regions of the visual cortex and determine how visual stimuli are processed there. Dissatisfaction with current widely used data analysis tools led laboratory members to design their own software in order to analyze fMRI experimental data with more accuracy. The software, which produces inflated and flattened cortical maps, allows more precision in identifying the location of brain activation. In addition, while more traditional studies typically use the ‘block design’ experimental paradigm (i.e., experiments are performed by using extended periods of ‘on’ versus ‘off’ activations), the software designed by the laboratory members is primed for detecting transient and rapid changes in the brain. The technique is becoming widely used, especially in research dedicated to visual cortex mapping. Yet, such novel ways for representing brain structure and its function have to be negotiated with the already well-established methods of fMRI brain representations.

3. MAPPING THE VISUAL CORTEX

When scientists talk about ‘mapping the human brain’ and ‘localizing functions’ in specific brain locations they are referring primarily to the cerebral cortex, the outer structure of the brain. The human visual cortex consists of multiple visual areas located in the posterior part of each hemisphere. Once the information arrives from the eyes, via visual pathways, to visual centers of the brain, it is passed from the early visual cortex to the higher order visual areas. The early visual areas tend to encode more elementary features, such as lines, whereas higher order visual areas contain neuronal groups that encode more complex features, such as edges, curves, and composition of features.

Retinotopic mapping is a procedure for mapping visual scenes onto the brain cortex. The idea is that the cerebral cortex contains a representation of the visual field that preserves some aspects of similarity with it. One can speak of a retinotopic map when adjacent locations in the visual field are represented in the adjacent locations in the cortex. The process of identifying retinotopic maps on the cortical surface involves the presentation of a patterned stimulus moving through the field of view of a subject being scanned. Because of the temporal match between the stimulus and the neuronal response, scientists are able to identify which parts of the visual cortex processes stimuli located at specific points in the visual scene. This allows scientists to assess the location and borders of a specific brain area. "In daily life, one would normally begin exploring a new geographical region only after first consulting the most accurate available map of that region. Analogously, to make progress exploring human visual cortex, we first need to have an accurate map of the areas which exist there" (Tootell et al., 1998: 175).

4. HOW TEXT AND IMAGES CONVEY THE EXISTENCE OF RETINOTOPIC MAPS IN HIGHER ORDER VISUAL BRAIN AREAS

The research article under examination suggests that, in addition to the well known retinotopic areas that exist in the early human visual cortex, there are retinotopic maps in the higher-order visual areas as well. The article reports findings of a map, located at the border between the visual and somatosenory cortices, which represents the angle of a remembered target[6]. The meaning of the text, and especially the results of the study, are in a significant manner produced through the brain images reproduced in the article.

The introductory part of the article lays out the historical background. It gives the overview of other studies and discusses methods that had the goal of defining maps on the cortical surface[7]. The text also points out to a lack of prior evidence about the particular topic of the paper. Next, the focus shifts to a description of the participants and the task of the study (illustrated in the Figure 1 of the article, not reproduced here). The role of this part of the text is to portray the precisely ordered steps of procedure and the technical equipment used[8].

Similar to the introduction, the final section of the article is mainly text-oriented. The section is dedicated to placing results in the broader picture: connecting it with larger scientific knowledge, predicting future research and attempting to suggest some interesting links between this study and other domains of cognitive neuroscience (i.e., vision and language).

---

[6] This is accomplished by asking participants to move their gaze over the visual field to the location where previously shown, and now remembered target (a small white square on a black background) was presented.

[7] Latour's (1987) gathering resources and allies to make dissent impossible.

[8] For a discusses on distinction between the normative sequence of project's course described in the method section of a paper and the actual performance of actions by members in the project see Lynch, 1985: 57.

Yet, importantly, the main part of the text, where the results of the study are presented, is heavily based on the visual material[9]. It opens by directly indicating the figure (Figure 2-1). The figure *shows*[10] the regions in the right hemisphere of one person[11] that were activated in a periodic fashion by the phase-encoded delayed saccade[12] task.

The figure is a complex image representing the right hemisphere superior parietal cortex from five different views: lateral, superior, posterior, medial-posterior and medial. By providing different views all at once, the figure allows for the viewer to see in one glance what would require movement through space and time if she were attempting to view an actual brain. Furthermore, the series of images works through cross-reference: each single image acquires its meaning by referring to the other images that are part of the Figure 2-1. For example, every single view on superior parietal cortex is replicated in 'folded' and 'unfolded' ways. So what is, for example, the Superior Temporal Sulcus on the representation of the lateral unfolded cortex can be explained by pointing at the representation of the same structure at the folded representation and vice versa.

---

[9] If we consider only the central part of the text, where the results are presented, we find that almost one half of sentences directly refer to figures.

[10] The image presents itself as having the role of an intermediary between the text and reality. Its function is to show the results of the study and to indicate real world phenomena. While the article's text may be associated with the author's opinion, i.e. the *question of believing*, images refer to what the text is about, i.e. the *question of seeing* (Latour, 1987: 48).

[11] Note how the images representing the brain activations in one person are used to talk about 'the general brain'. For an account on the 're-appearance' of the experimental participant in the fMRI data and her subsequent erasure from it, see Alač (submitted).

[12] Saccades are mall rapid movements of the eye.

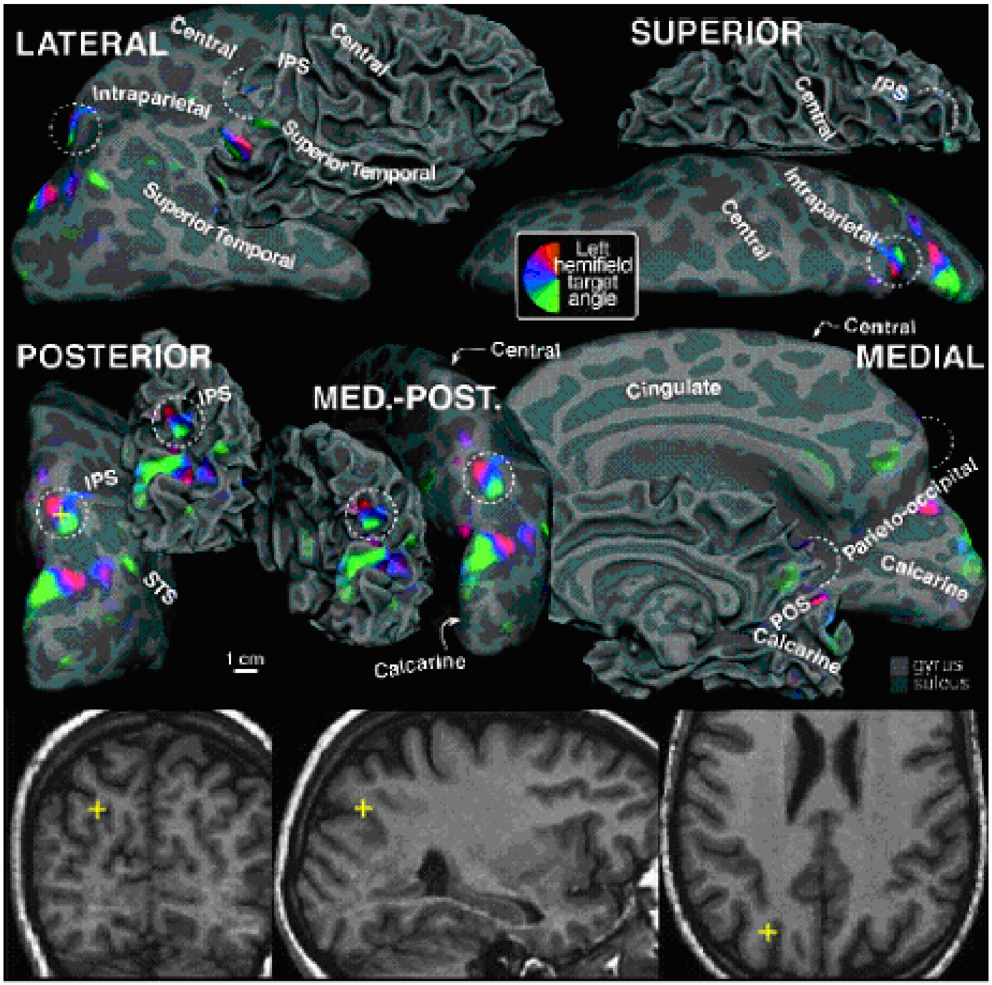


FIGURE 2-1: Figure 2 from Sereno et al., 2001

But why are we presented with these parallel representations, when the unfolded ones look quite ‘unnatural’ to the unaccustomed viewer? The representation of a 3-D folded cortex preserves the more ‘natural’ appearance of the cortex; however, this type of representation does not show what is buried in the sulci (or grooves) on the surface of the brain. Because of the high percentage of the buried structures of the cortex, ‘distance measured in 3-D space between two points on the cortical surface will substantially underestimate the true distance along the cortical sheet, particularly in cases where the points lie on different banks of a sulcus’ (Fischl et al., 1999: 273). Hence, the unfolded view (a feature produced by the software used in the PI’s laboratory), even though more remote from the actual appearance of the real brain, allows more precise measures of the cortex. By showing and not hiding, the software used in the PI’s laboratory produces views of the brain that are particularly effective for scientific applications such as brain mapping.

This type of visual representation highlights the direct interdependence between ‘seeing’ and ‘knowing’. If scientists were only interested in statistical measurements, the flattened brain representations would be totally superfluous. Why create unnatural representations if the possibility to perform such measurements is not enhanced? The flattened representations, however, enable scientists to gain knowledge through seeing. The transformation of the folded cortex representation into the flattened cortex representation enhances the visibility of the object and allows scientists to exercise more control over it.

But what Figure 2-1 specifically directs our attention to are ‘maps of remembered locations’ (and flattened brain representations enhance such direction of attention). The image not only makes directly visible the structure of the brain, but also inscribes maps generated through recording action in the brain. The fMRI technique thus makes visible and spatial what is otherwise invisible and temporal. In Figure 2-1 these functional maps are represented through brightly coloured patches (red, blue and green) on gray-coloured representations of the cortex. Bright colours show which part of the human brain is active when processing contralateral remembered targets (red for upper left visual field; blue for the left horizontal visual field, and green for the lower left visual field)[13] [14].

Thus, the colours used in Figure 2-1 can be sorted in 3 groups with respect to what they are associated with: 1) the background, 2) structural images of the cortex, and 3) maps of remembered targets in the parietal cortex. The background, whose role is only to contrast that which is of interest, and has no intrinsic interest on its own, is black, erasing in such a way, the

---

[13] Note that, since the choice of colour is not standardized, but depends on the authors of the scientific report, the text/figure has to provide an explanation of what different colours stand for.

[14] The correspondences can be calculated because of the temporal match between the stimuli and the neuronal response. Since time is utilized in such manner, colours can code different phases of visual processing. This allows for the identification of the corresponding location in the visual field, and its representation in the visual cortex.

body of the person being scanned [15]. The structural images that represent what is static, and help us situate what is of interest, are in tones of grey. In contrast, the maps representing cognitive activity, which are meant to be the focus of attention, are in bright colours. By rendering them in bright shades, certain selected and highlighted portions of the image are marked as holding special meaning. Notice also how various types of graphical signs, that function as tools for selection, are inscribed on the image. Text labels, which route our comprehension of what is represented, are superimposed on the images; zones of interest are indicated by the dotted circles and yellow crosses; legends for colours and scale are given. In this way, like text, visual representation is constructed so as to rule out any interpretation not intended by the author (Bastide, 1990).

Yet, despite all the guidance through the figure, some brain representations that are part of Figure 2-1 strike the observer as quite different from the others. The basic distinction is between the more standard structural images that report the statistical data of 12 subjects, and those of one subject's brain produced with the program for data analysis designed by the PI and his collaborators. The bottom part of the figure represents three brain slices (coronal, sagital and axial) given in the structural form. Through a process of translation, the centre of what are the parietal maps on the images of the cortex is inscribed with a yellow cross on the structural images. The location is also given in Talairach coordinates[16]. Because the larger scientific community uses the Talairach normalization approach, in addition to the conventions

---

[15] For a similar discussion on ultra-sound imaging see Hartouni (1991).

[16] When an individual brain is imaged, mapped and the data of interest are measured, the representations produced aren't yet necessarily meaningful to the scientific community. Because of the idiosyncrasies of every brain, images of a single brain often have to be matched to a standard. In order to be shareable with other scientists they have to match what is called the ‘standard brain’. In many laboratories, the comparison of data across subjects is based on the Talairach's (Talairach et al. 1967; Talairach & Tournoux, 1988) 3-D normalization approach: representations of single brain are projected, through the registration procedure, to ‘Talairach brain’. Hence, it can be said that the representations acquire meaning with respect to that standard. In other words, ‘Talairach’ is a historically shaped physical and conceptual construct for production of meaning that, by selecting some relevant elements, provides a form of reference shareable with other scientists.

adopted by the laboratory members (exhibited in the upper parts of the Figure 2-1), the results of the study are presented in this form as well. The PI and his co-workers note significant drawbacks in the Talairach approach, which is based on 3-D stereotaxic coordinates rather than on position relative to the 2-D cortical sheet (e.g., Fischl et al., ibid). Nevertheless, they translated the results of the study into this system because it is the format accepted and used by the cognitive neuroscience community at large. In such a way, the figure inscribes in its form the coordination between the standards and conventions adopted by the fMRI scientific community at large, and the visual presentation formats employed by the laboratory members. In order to find out more about standards, conventions, and their negotiations, I'll focus next on the article's publication process.

5. REVIEW PROCESS

Publication of an experimental research article is preceded by a stage where the idea and the design of the experiment are assembled, a long period of meticulous work in the laboratory, and the activity of writing and revising. In addition, a scientific text is also the result of the various revisions based on judgments of journal reviewers' before its actual publication. While the former stages tend to leave meagre traces in the article's texture (e.g., Knorr-Cetina, 1981; Lynch, 1983), the later stages are usually completely invisible (except for rare occasions when anonymous reviewers contributions are mentioned in the article's acknowledgements section). By bringing this often 'black-boxed' dimension of the manuscript's production into the picture, I will illustrate the role that the visual brain representations played in the practice of scientific fact production.

5.1 WHAT HAPPENED BETWEEN THE MANUSCRIPT'S INITIAL DENIAL AND SUBSEQUENT ACCEPTANCE FOR PUBLICATION?[17]

One month after the first submission of the paper to the journal, the PI and his colleagues received a letter from the editor. The letter informs the authors of a negative decision regarding the publication of the manuscript. However, it also states that it may be possible for the authors to resubmit a revised version of their manuscript after completing additional experiments and analyses following the referees' suggestions. The paper is resubmitted after approximately seven months, and is finally accepted for publication one month after the resubmission. The paper is published during the following winter, about a year after the first submission. But what happened between the first letter from the editor and the second one, which informs the authors of the paper's approval for publication? To tackle this question I will analyze the PI's second letter to the editor accompanying the resubmission of the manuscript.

In the letter, the PI states that, following the reviewers comments on the previous version of the paper, new data were collected and new task variations were deployed. It is explained that the paper has been rewritten, and, most important for my analysis, all figures have been redone.

The letter reveals that the reviewers were dissatisfied with one of the tasks that the participants performed while in the scanner. They also express dissatisfaction with the way in which statistical significance was visually represented. In addition, one reviewer wondered if the spatial maps were the same across different conditions and different scanning days. An accidental reversal in the colour scale[18] was noticed as well. And, importantly, the reviewers expressed their concern about the activations of the earlier visual areas. These early brain

---

[17] I would like to thank the participants in this study for their generosity, and at the same time to point out that every claim present in this text reflects only my own views.

[18] Red, instead of being used to represent the upper left visual field, was used to represent the lower left visual field; likewise, green instead for the lower left visual filed, is used for the upper left visual field.

activations are shown in Figure 2-2a which accompanied the original manuscript submission; they are however missing in Figure 2-2b which is a part of the published article.

The PI's response is divided into five sections: 'Quantitative Measures of Response Magnitude', 'Reliability of Results', 'Task Adequacy', 'Unexplained Activity in Earlier Areas', and 'Small Changes'. Four parts directly deal with the figures. Only the section on the addition of the new experimental task can be seen as not directly dependent on the visual representations.

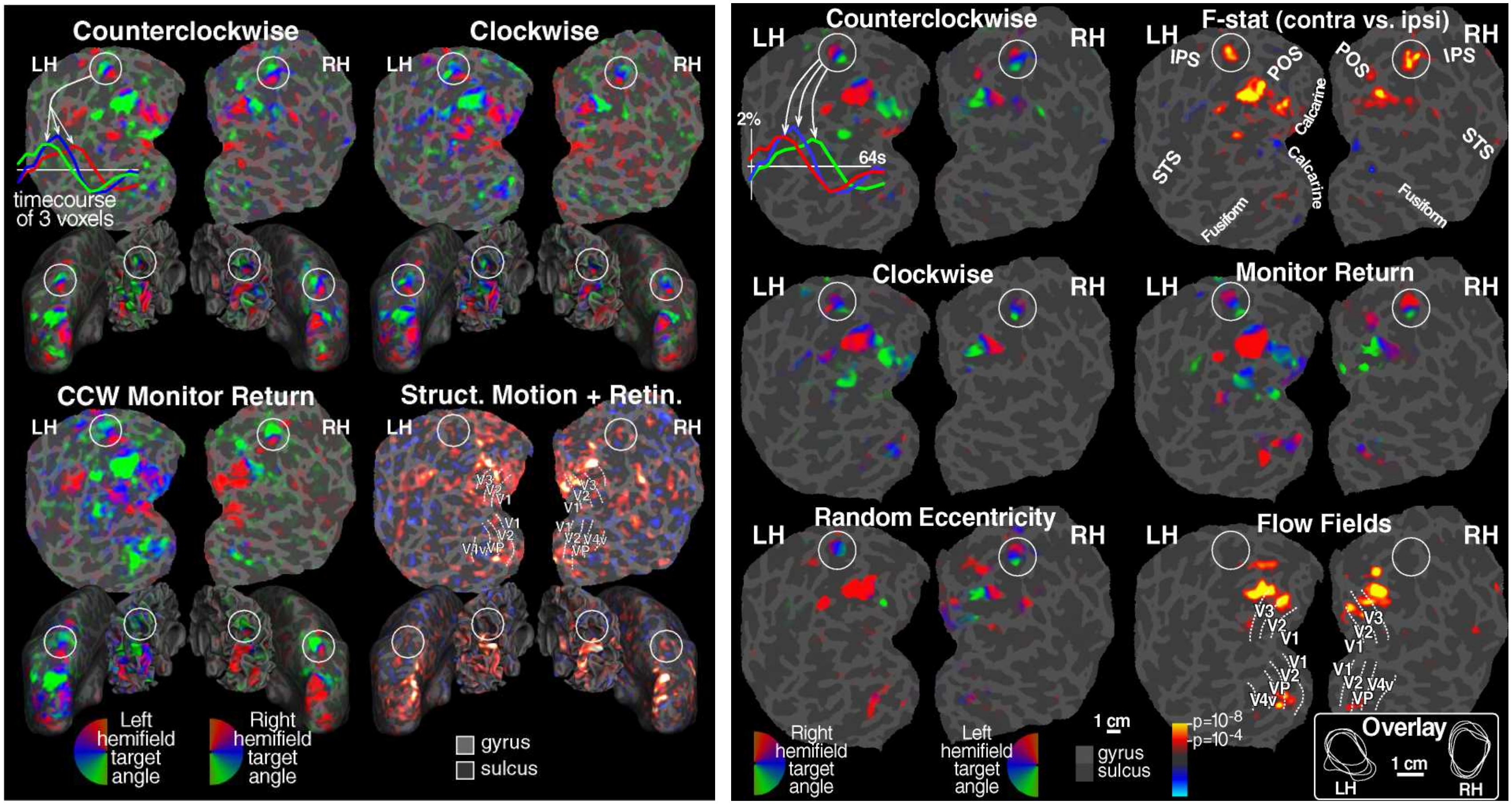


FIGURE 2a: Original Figure3 - FIGURE 2b: Published Figure3

FIGURE 2-2: Figure 3 from Sereno et al., 2001

5.1.1. INCLUSION OF A NEW TASK

The inclusion of a new experimental task is one of the substantial modifications that appear in the final version of the paper. One of the reviewers (Reviewer 3) was dissatisfied with the basic delayed saccade task. In this task, the participant fixates at the centre of the image while a peripheral target (a single dot) appears at 15° of eccentricity. Shortly after, a ring of blinking distractors (multiple dots) appears on the screen. Next, the participant has to move her eyes to the remembered location and back. The task appeared too predictable to the reviewer.

To respond to the reviewer's criticism, the researchers added task where the eccentricity of the next target is unpredictable (i.e., the target can appear in the range of 1° to 15° of eccentricity), and they reported new data, which are very similar to the data from the original task. This modification in the experimental design is apparent in the published figure (Figure 2-2b). In comparison with the original version (Figure 2-2a), the published figure contains a representation of the extra condition (lower left part of the Figure 2-2b). The modification, even though not directly dependent on the visual representation, is reflected in it. Yet, all the other modifications, which I will next discuss, were achieved by directly modifying the original figures.

5.1.2. NEW ANNOTATIONS

In addition to the representation of the new experimental condition, the published figure (Figure 2-2b) contains another brain representation not present in the original figure: the representation of statistical significance in the upper right hand-side of the figure. As already noted, the authors of the article proposed an innovative way of visually representing brain activation. They used colour to present at the same time the statistical significance and the existence of the map (see, for example, the left hand-side of the figure 2-2b). One of the reviewers (Reviewer 2), however, was unsatisfied with the way in which the colour saturation represented the statistical significance: ‘Reviewer 2 suggested that the representation of the statistical significance was *imprecise and difficult to quantify* […]’ (second letter to the editor),(italics are mine). The reviewer’s comment shows that colours are perceived as an inadequate way to represent the reliability of the measure of levels of neural activity in the cerebral cortex. It is believed that the quality of representation should be strengthened by allowing for precision and ease of quantification.

In order to answer this critique, the PI and colleagues inserted a significance value (‘F-statistic’) representation (Figure 2-2b, upper right side). If comparing the original submission and the figure in the published version of the article, one can notice that the data shown in the representation on the left hand-side of the top row (Figure 2-2b), where colour is used to represent both the significance (colour saturation) and the existence of a map (different colours), are re-represented on the right side of the figure, so that the brain representation contains only a significance value. The new plot allows the viewer to read by comparing the colour plots in more traditional terms already established by the scientific community. The PI writes: ‘This makes it easy to *verify* that colour saturation is indeed an effective way to portray significance while at the same time demonstrating the existence of the map’ (italics are mine).

The PI’s response is a clear example of scientific negotiation between more traditional, widely accepted forms of experimental data presentation, i.e., p-values, and the newer ways of representation, characteristic of the practice of retinotopic map identification and largely used by the PI and his group. The mediation is achieved by the inclusion of the more widely accepted form next to the original plots. This is done in order to make the comparison possible and to allow for verification. The reviewer was not completely satisfied because the colours didn’t allow for an adequate quantification. The authors provided the additional notation that enabled verification, and hence, aid in the understanding of the image where colour is used to indicate polar angle.

Other annotations are added to the single-voxel time course plot (in the upper left-hand corner of the image). The time course plot functions as a value judgment on the brain representations. It indicates the quality of the fMRI signal (i.e., the technique for generating retinotopic maps produces activation of a portion of a visual field map, that is, parts of the parietal cortex map that represent different portions of the visual field respond at different points in time), and hence emphasizes the clear borders and the well-defined appearance of the cortical map. The cleanliness and the visibility of the data, usually not present when traditional methods are applied (i.e., when software similar to the one designed by the laboratory members that records responses to phase-encoded retinal stimulation are not applied), is emphasized. The plot points out that the phase-encoded paradigm, used by the researchers, produces less ‘noisy’ data, where different phases of activity in nearby voxels can be distinguished: ‘Here, we were able to show under the more stringent conditions of a phase-encoded mapping paradigm that the parietal cortex map is in fact quite visible, even at the single voxel level’. In order to attain further specifications, the annotation ‘2%’ is added to the y-coordinate, and the annotation ‘64s’ is added to the x-coordinate of the time course plot. By doing so, the value judgment expressed by the plot is rendered quantitatively more precise.

Another new feature of the published image is a graphical representation of the overlay situated in the lower right-hand side of the image (Figure 2-2b). The goal of the representation is to indicate the way in which different maps were overlaid on top of each other. This overlaying of representations is part of the answer to the criticism that the spatial maps obtained from the experiment were not identical across different experimental conditions and different scanning days. The authors believe that this type of remarks depends on the fact that the original representations make a direct comparison between conditions too difficult. Hence, in order to reduce this difficulty in comparing different representations, scale bars with significance values, as well as the overlaying of images were added. In such a way, the representation of the overlay is again at once a comment on different brain maps and a guide for looking at such maps. It encourages the comparison between brain maps obtained for different conditions and across different scanning days. The authors also pointed out that the overlay of images of the same maps acquired across different sessions is more precise in comparison to what is usually presented in imaging papers published in the same journal.

It is also immediately noticeable that the new version of the Figure 2-2b does not contain folded and inflated brain representations. This omission is a consequence of one reviewer's complaint that the multiple representations of the same data set in the original image were confusing. This modification determined the appearance of the Figure 2-1 (analyzed in the first part of my text). Because the authors had to omit folded and inflated brain representations from Figure 2-2b, the representation located in the upper left-hand side of Figure 2-2b, converted into a larger scale view in inflated and folded forms, is represented in Figure 2-1. Thus, Figure 2-1 preserves the less-conventional type of representation. According to the authors, this type of representation provides a richer knowledge about the organization of the cortical map, while a 'less confusing' and simpler representation of the data is produced in Figure 2-2b.

In the 'Small Changes' section of the PI's letter to the editor, the authors state that they adjusted the accidental reversal of colours in the brain representations that one reviewer noticed (i.e., instead of using red to represent the left upper visual field, green was used, and vice versa). The fact that the reviewer noticed the reversal clearly indicates that she/he was paying close attention to the images.

5.1.3. ACTIVITY IN EARLIER VISUAL AREAS

The most striking feature of the new figure is the different colour distribution on the brain images. This aspect of the figure directly touches upon the problem of localization of function. In the second letter to editor, the PI writes:

> In our previous submission, we rendered figures with the contrast on the colour saturation function set too low – especially for the printed page. In the new version we have increased the contrast so that the viewer is not distracted by regions of low significance. Many of the complaints of Reviewer

> 2 and 3 centered on regions of activity beyond the two that we focus on – which were, the putative human LIP and a more posterior region overlapping V3A. The higher contrast rendering make it easier to distinguish these two regions of high significance and interest.

This passage presents different points of interest. First, it shows how manipulation of the image is an important aspect of meaning production. If the two images in Figure 2-2 are compared, it is easy to notice that, while the image from the original submission (2a) contains colour patches all over the cortical surface representation, the new, published image (2b), contains colour only at a few, well-defined locations. As the passage states, this change is achieved by raising the threshold level. In such a way, the figure displays only the most statistically significant values, while the others are veiled away. This remedy allows for the more efficient management of the viewer's gaze over the image: 'the viewer is not distracted by regions of low significance'.

The change in threshold rising is monotonic, which means that there is never a case where less significant data become represented as brighter than the more significant one. Yet, importantly, the scientific claim and its presentation are made more apt by working directly on visual representations of the experimental data, while keeping the quantitative data constant. This manipulation of visual representations is directly involved in the process of changing the criteria for what is relevant, and what is not; for what has to be shown and explained, and what does not.

The rising of the threshold and the modification of the visual brain representation also speak about expectations held by viewers familiar with the fMRI culture. In order to ground the mind in the brain, cognitive neuroscience is accustomed to assigning specific cognitive processes to particular areas of the brain. To find the neuro-anatomical substrates of mental activity, cognitive processes are correlated with well-defined locations in the brain. The reviewers expected to see specific and unique 'spots' on the brain surface representations that would univocally and clearly prove the point that the higher level visual areas present retinotopical mappings. The other, smaller effects, representing attention effects in earlier visual areas, are statistically significant, and their existence is scientifically proven (e.g., Gandhi et al., 1999). However, as the PI pointed out during a conversation, ‘people like less spots’; the existence of different ‘spots’ all around the brain representations ‘complicated the story’ and therefore, it was wiped out by raising the threshold. Hence, what the published image represents are only restricted and well-delimited brain areas that responded in the most statistically significant way to the given stimuli.

It is also interesting to note that the difference between the printed page and its digital form is pointed out: ‘In our previous submission, we rendered figures with the contrast on the colour saturation function set too low – *especially for the printed page*’. The images sent to the journal in digital form became darker when printed on paper. This unexpected accident caused the contrast between the different colors, and visibility and differentiation between different signifiers for brain areas and meaningless units (‘low-significance regions’), to get lost. The representational form, and its material support play an active role in the process of meaning construction.

6. INTERACTION OF VISUAL AND QUANTITATIVE ASPECTS OF FMRI REPRESENTATIONS

> The question regarding the role brain images play in the production of knowledge is at the core of recent research on fMRI and PET. Anne Beaulieu (2002) points out that scientists who use fMRI and PET reject the importance of brain representations as visual data. Beaulieu’s interview data show that, scientists highlight the potential of brain imaging measurement to render spatial components and anatomical referents. At the same time they downplay the visual form this information takes in order to emphasize the quantitative information it represents.

This negation of the importance of visual knowledge in brain mapping research is understood by Beaulieu as related to the way evidence is evaluated in modern Western science. Beaulieu argues that because it has been regarded as appealing first to the senses, as opposed to reason, and hence is seen as lacking a solid relationship to the truth, visual evidence is judged as not having a particularly high position in the hierarchy of types of scientific evidence. In accordance with this claim, Beaulieu's interviews with brain mapping researchers show that researchers affirm that their practice consist of *measurements* in the brain, and are not directed towards obtaining *images* of the brain. This points toward a historical divide between visual and quantitative modes of representation. The interviewees claim that those most interested in the visual aspects of brain mapping techniques are usually clinicians, not scientists, suggesting a hierarchy in which the visual is associated with the lower echelon of applied research.

In the present analysis, the reviewers' comments and their leaning toward the quantitatively oriented conventions for data presentation echoes the stand expressed by Beaulieu's interviewees. Yet, the analysis of the publishing process also indicates that the divide between visuo-spatial and quantitative modes of representation in brain mapping research, suggested by Beaulieu's interviewees, is debatable.

The description of the negotiation practices that took place during the manuscript's revision phase largely illustrated modifications done on visual representations. During this phase, complex interrelationships between quantitatively oriented conventions for data presentation, and visual representations came into play. These negotiations, however, cannot be characterized as interactions marked by an irresolvable tension between visuo-spatial and purely quantitative representations. Quite the contrary; reviewers' comments reveal that they were primarily interested in what was conveyed through the images. Reviewers didn't seem particularly concerned with the way in which the data were presented. Instead, they spoke about the facts proposed by the article (i.e., the existence of retinotopical maps in the higher order visual areas). Their knowledge of these facts came substantially from the manuscript's figures. The reviewers questioned elements (such as the appearance of multiple regions of cortical activity) that were apparent in the figures, not in the text. Along the same lines, the figures were a central target of the manuscript's modification process. This work, largely directed towards the figures' modifications, resulted into the article's acceptance for publication.

7. SEEING AS NEGOTIATION

The fundamental co-dependence between quantitative and visuo-spatial modes of representation can be associated to the fact that fMRI brain images are at once visual and quantitative. They are visual representations of quantitative information. In other words, they are *data graphics* that display measured quantities largely by means of colour. Speaking of data graphics, Edward Tufte writes:

> Often the most effective way to describe, explore, and summarize a set of numbers – even a large set – is to look at pictures of those numbers. Furthermore, of all methods for analyzing and communicating statistical information, well-designed data graphics are usually the simplest and at the same time the most powerful (Tufte, 1983:9).

Tufte's suggestion that the 'pictures of numbers' have a role in *describing* and *communicating*, as well as in *exploring* and *analyzing* is related to various proposals in cognitive science that deal with the relationship between external, visual representations and cognition.

Since the pioneering work of Larkin & Simon (1987), students of *diagrammatic reasoning* have pointed out that the way in which a problem is represented - whether, for example, it explicitly preserves the information about the topological and geometric relations among the components of the problem - is crucial for our performance in problem solving (for reviews see Glasgow, Narayanan, Chandrasekaran, ibid; Anderson, Meyer, Olivier, ibid). Similarly, researchers have pointed out the role of *re-coding* in cognitive task performance. In re-coding an abstract mathematical problem is re-presented into a form which exploits the specific computational powers of our visual system (Kirsh, 1995). The idea has also been called *representational re-description* (Karmiloff-Smith, 1992) and *trading spaces* (Clark &

Thornton, 1997). Moreover, proponents of cognitive semantics and embodiment (e.g., Lakoff, 1987; Fauconnier & Turner, 2003) have pointed out that understanding of abstract entities and processes is often based upon biological capacities and experiences of functioning in a physical and social environment. One concept from the research domain is the idea of *conceptual scaling* (Fauconnier & Turner, ibid.). Through processes of conceptual scaling, what is invisible or not visual (as in the case of quantitative data) is transformed into what we can directly experience (i.e., visual representations). A concept is on *human scale* with respect to the ease with which it is produced, manipulated, or remembered. Since humans use vision as our primary sense, it is cognitively more apt for us to analyze, manipulate, and describe complex and abstract entities as visual and spatial.

Yet, to explain the fact that visualization deeply shapes scientific activity in terms of the processes related to individual organisms does not entirely do justice to the reality that emerges from the revision processes described in this text. The preceding analysis emphasizes the centrality of the brain images in the production of scientific fact. Yet it also illustrates the inevitable embeddedness of scientific text in a particular social, cultural, and historical context, as well as its dependence on the audience. These elements are more easily overlooked when one is immersed in the ‘clean’ world of experimental procedures. Yet they become strikingly real in scientists’ everyday practices, such as the revision process.

One of the symptomatic points that came up through the revision process was the reviewers’ dissatisfaction with the appearance of ‘regions of activity beyond the two that the manuscript focused on’. The reviewers’ remarks implied inclinations and expectations to see specific locations of activation, as is customary in fMRI literature. As this negotiation of the ‘number of cortical activation spots’ shows, the link between the image and its referent is always mediated through the filters of social and cultural norms and expectations. ‘Seeing’ of reality becomes ‘clearer’, ‘less confusing’, and ‘easier to verify’ when it follows the parameters of the scientific culture of which it is part.

ACKNOWLEDGMENTS

This work was supported by grant # 9873156 from the National Science Foundation (PIs Jim Hollan, Edwin Hutchins, David Kirsh). I would like to thank Anne Beaulieu, Lisa Cartwright, Seana Coulson, Hollie Crower, Trudy Dehue, Edwin Hutchins, Martin Sereno, Cecilia Ubilla, Hsin-Hao Yu, Ying Wu, and the participants from the Workshop on Objects of Objectivity for their contribution to this project.

WORK CITED

ALAC, M.
Submitted 'From trash to treasure: learning about the brain through multimodality'.

ALAC, M., & E. HUTCHINS
In press 'I see what you are saying: action as cognition in fMRI brain mapping practice'. *Journal of Culture and Cognition*, 4.3.

ANDERSON, M., MEYER, B., OLIVIER, P. (Eds.)
2002 *Diagrammatic Representation and Reasoning*, London: Springer-Verlag.

BASTIDE, F.
1990 'The iconography of scientific texts: principles of analysis', In M. Lynch & S. Woolgar, (eds.).

BAZERMAN, C.
1988 *Shaping Written Knowledge: The Genre and Activity of the Experimental Article in Science*, Madison: University of Wisconsin Press.

BEAULIEU, A.
2000 'The brain at the end of the rainbow: the promises of brain scans in the research field and in the media', In *Wild Science: Reading Feminism, Medicine, and the Media,* J. Marchessault & K. Sawchuk, eds., London and New York: Routledge.

2002 'Images are not the (only) truth: Brain mapping, visual knowledge, and iconoclasm', *Science; Technology, & Human Values*, 27, 53-86.

BOURDIEU, P.
1984 *Homo academicus*. Paris, Les editions de Minuit.

CLARK, A. & THORTON, C.
1997 'Trading Spaces: Computation, Representation, and the Limits of Uninformed Learning', *Behavioral and Brain Sciences* 20, 57-90.

FAUCONNIER, G. & TURNER, M.
2002 *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*, New York: Basic Books.

FISCHL, B., SERENO, M.I., TOOTELL, R. B.H., DALE, A.M.
1999 'High-resolution intersubject averaging and a coordinate system for the cortical surface', *Human Brain Mapping*, Volume 8, pp. 272-284

GANDHI, S.P., HEEGER, D.J., BOYNTON, G.M.
1999 'Spatial attention affects brain activity in human primary visual cortex', *Proc.Natl.Acad.Sci*, *USA*, Vol.96, pp.3312-3319.

GLASGOW, J., NARAYANAN, N. HARI,, CHANDRESEKARAN, B. (eds.).
1995 *Diagrammatic Reasoning: Cognitive and Computational Perspectives*, Cambridge, MA: MIT Press.

HARTOUNI, V.
1991 'Containing Women: Reproductive Discourse in the 1980s', In C. Penley and A. Ross (eds.), *Technocultur*e, Minneapolis: U of Minnesota Press.

KARMILOFF-SMITH, A.
1992 *Beyond Modularity*, Cambridge, MA: MIT Press.

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

KNORR-CETINA, K.
1981 *The manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science*, Oxford: Pergamon.

LAKOFF, G.
1987 *Woman, Fire, and Dangerous Things*, Chicago: Chicago UP.

LARKIN, J. & SIMON, H.
1987 "Why a diagram is (sometimes) worth ten thousand words", *Cognitive Science*, 11:65-99.

LATOUR, B.
1987 *Science in Action*, Cambridge, MA: Harvard UP.

LYNCH, M.
1985 *Art and Artifact in Laboratory Science: A Study of shop work and shop talk in a research laboratory*, London: Routledge & Kegan Paul.
1990 "The Externalized Retina: Selection and mathematization in the visual documentation of objects in the life sciences", In M. Lynch & S. Woolgar (eds.).
1993 Scientific Practice and Ordinary Action: Ethnomethodology and social studies of science, Cambridge: Cambridge UP.

LYNCH, M. & S. WOOLGAR (eds.)
1990 *Representation in Scientific Practice*, Cambridge, MA: MIT Press.

SERENO, M.I., S. PITZALIS, A. MARTINEZ
2001 Mapping of controlateral space in retinotopic coordinates by a parietal cortical area in humans. Science, 294:1350-1354.

TALAIRACH J., SZIKLA G., TOURNOUX P., PROSSALENTIS A., BORDAS-FERRER M., COVELLO L.
1967 Atlas *d'Anatomie Stéréotaxique du Telencephale*, Paris: Masson.

TALAIRACH J., TROUNOUX P.
1988 *Co-planar Stereotaxic Atlas of the Human Brain*, Stuttgart: Thieme.

TOOTELL, R.B.H., HADJIKHANI, N.K., VANDUFFEL, A.K.L., MENDOLA, J.D., SERENO, M.I., DALE, A.D.
1998 'Functional analysis of primary visual cortex (V1) in humans', *Proceedings of National Academy of Science, USA*, Vol.95: 811-871, February.

TUFTE, E.R.
1983 *The Visual Display of Quantitative Information*, Cheshire, Conn.,: Graphics Press.

This chapter is reprint of

Morana Alač: "Negotiating Pictures of Numbers", *Journal of Social Epistemology*, 18:2, 199-214, 2004.

CHAPTER 3

FROM TRASH TO TREASURE: LEARNING ABOUT BRAIN IMAGES THROUGH MULTIMODALITY

ABSTRACT

Cognitive neuroscientists use functional magnetic resonance imaging (fMRI) technology to generate digital images of the human brain. An fMRI image, as a final product of the scientific work, does not document movements and sounds that were present when such an image was recorded. Yet, a focus on actual moments of scientific practice reveals that such forgotten elements of practice can play important roles in understanding and knowledge acquisition. The multimodal interaction among scientists and digital screens shows how movements of the experimental subject and the scanner noise are performed to make images meaningful. Moreover, it suggests that the phenomena whose detection is crucial for a scientific reading of the brain images, such as motion artifacts, become visible as a result of coordination of various semiotic modalities (i.e., images, talk, body movements, gesture, etc.).

1. INTRODUCTION

Cognitive neuroscientists use Magnetic Resonance Imaging (MRI) technology to produce images of brain structure as well as its function (fMRI). Functional MRI is used to visualize the local changes in magnetic field properties occurring in the brain as a result of changes in blood oxygenation. An fMRI image can show the degree of activity in different areas of the brain; if the image is obtained while an experimental participant is engaged in a particular cognitive task, the image can indicate which parts of the brain are most active in that task[19]. In the past decade, a revolution in brain-imaging technology has made such images omnipresent not only in scientific circles, but also in popular media (Dumit, 2004[20]).

It is commonly assumed that the power of the fMRI technology resides in its capacity to produce knowledge about the brain and cognition through vision (Kutas *et al*., in preparation). By observing fMRI images scientists can *see*, and hence *comprehend* the workings of the human brain and mind[21].

Rather than simply accepting the assumption, I'll question the capacity of fMRI images to expose the human brain through an "abstract sense of vision" (Lenoir, 2004: XXII; see also Alač & Hutchins, 2004). To do so, I'll ask: How is the *seeing* achieved? How do young practitioners acquire knowledge about brain images? And, in particular, what is the relationship between such a capacity (i.e., the capacity to directly see and understand), and publicly shared, embodied practices? For example, can interpretation of the brain images, in addition to vision, concern hearing or body movement?

Studies of science and technology have often pointed out (Lynch, 1985; Latour, & Wollgar, 1979; Knorr-Cetina, 1981; Star, 1989) that final accounts of the scientific work (i.e., in form of scientific articles, conference talks, or newspaper reports) inform about general truths, while standardizing, averaging and filtering out local contingency or individual differences.

---

[19] For an excellent introduction from socio-anthropological point of view, see Roepstorff, 2004.

[20] Dumit's work concerns *Positron Emission Tomography* (PET) images. However, his observation can be straightforwardly applied to fMRI.

[21] For a discussion on importance of visual aspects of fMRI see Alač, 2004.

“One of the mandates of science is to create generalizable results, which are meant to be universal, and this mandate is often conflated with the deletion of local contingency” (Star, ibid.: 93). “The published data reveal, rather than hint at; articles state, rather than guess at; and subjects line up docilely to be counted. Their attempts to run away, their individual differences, and their resistance to the experiments are erased from the public record.” (ibid.: 65) In fact, with the exception of single subject studies, method sections of fMRI articles list repeatable and well defined steps of the experimental procedure, while reporting on experimental participants in terms of the number of subjects that were involved in the study, their gender, and social group. In such a way they make individual bodies utterly unavailable. In other words, “the published images talk about “*the* brain, not *a* brain.” (ibid.: 90)

Yet, by looking *beyond* and *before* these public records, I wonder what is the function of the individual and particular in the situated use of unformulated practical competences. Do the bodies of single subjects play any role in the actual practice of fMRI? Do practitioners, to acquire and understand principles of fMRI data analysis, enact behavior of individual bodies? Can such events tell us something about the way in which researchers understand central aspects of their practice?

To tackle these questions, I study brain imaging research in its every-day context of practice. I analyse events of apprenticeship wherein young practitioners learn through action: they are trained to read fMRI brain representations by actively participating in data analysis sessions. My approach is grounded in *practice-based theory of knowledge and action* (see Goodwin, 1994: 606). My aim is to reconcile questions of interest to *science and technology studies* with observations of micro-aspects of interaction. To do so, I use insights from *distributed cognition* (Hutchins, 1995) and *conceptual integration theory* (Fauconnier and Turner, 1998, 2002). In order to examine fMRI practice, I combine traditional long-term participant observation with the micro-analysis of specific occurrences of events and activities (e.g., Goodwin, 1994, 1995, 2000, 2003; Heat & Hindmarsh, 2000; Hutchins & Palen, 1997, Suchman, 2000). To find out how scientists work with and understand fMRI images, I perform micro-analysis on digital video recordings of scientific practices. I use these recordings to identify and track the semiotic structures that scientists bring into coordination with the images in the process of knowledge construction. I utilize the findings from my long-term ethnography to provide warrants for my analytic judgments concerning how things and events are made relevant to one another in a minute instance of scientific practice.

The findings reported here are part of a study conducted in three laboratories that employ fMRI technology to study the brain and cognition. The study lasted twelve months and involved fifteen participants. In addition to video recordings, data collection methods included direct observation, semi-structured interviews, and analysis of documents, such as research articles, laboratory manuals, and scientific correspondence.

2. MOVEMENT AND NOISE AS OBSTACLES FOR fMRI RESEARCH

Upon being asked for what is crucial in obtaining good experimental data, one of the senior researchers involved in the study replied “a good subject,” that is, “one who doesn’t move during the scanning session.” This statement is important not only because it attributes centrality to the experimental subject, but also because it stresses how harmful a subject’s movement is. In other words, fMRI researchers consider the *artifact*[22] associated with the subject motion to be one of the greatest obstacles to successful data collection and analysis.

Since even small movements of the tongue or jaw can cause artifacts in the data by altering the local magnetic field around scanned areas, the head movement of the subject has to be controlled as thoroughly as possible. In order to control movement, different head-restraint systems, such as thermal plastic masks, vacuum pillows, and bite-bars are used. Figure 3-1 shows an experimental subject just before entering the bore of the MRI scanner[23]. The subject lies supine on the scanner table while her head is restrained with a thermal plastic

---

[22] For an explanation and definition of research artifact see Lynch, 1985: 81-83.

[23] The fMRI scanner is essentially a magnet containing gradient coils and the radio frequency coil to produce the magnetic field necessary for imaging.

mask. While warm, the mask takes on the shape of the subject's face. With cooling, it becomes rigid and hence prevents the subject's head from moving.

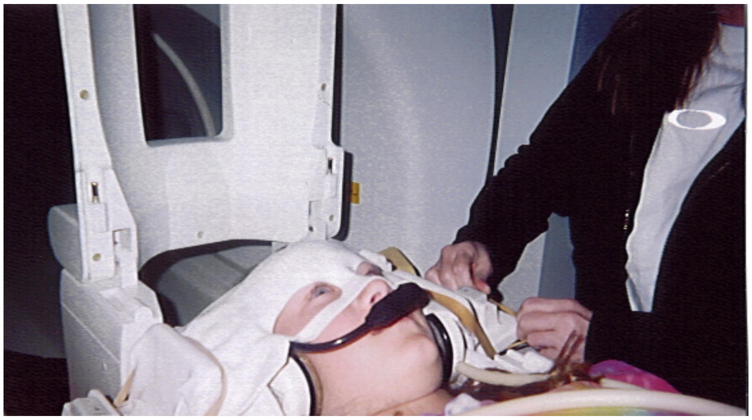


FIGURE 3-1: An application of the thermal plastic mask

During fMRI experiments subjects also wear headphones and ear plugs. Such devices are used for isolation from the noise that the scanner produces. Scientists are well aware that the high decibel acoustic noise generated by the scanner can impede effective research. Masking of background noise is especially important when acoustic, rather than visual stimuli are used (e.g., research on language comprehension).

In short, scientists strive to design and carry out experiments in which movement of the subject and scanner noise are reduced to a minimum; movements and noise are nothing more than obstacles for research.

For the purposes of this analysis, however, I will focus on these *obstacles*, rather than disregarding them. I will look at how such phenomena are reenacted as part of discursive social action and what their role is in knowledge production. I will show how these unwanted aspects of the scientific procedure are reused by practitioners and made functional. In order to trace ways in which such elements leak back into the process of scientific sense making, I will focus on traditionally less explored communication modalities, such as non-linguistic sounds, body movements, and gestures, I will describe how, in two instances of fMRI apprenticeship, such modalities are combined with brain images to produce understanding.

3. LISTEN TO UNDERSTAND

As already mentioned, fMRI is a technique associated with image production. Much effort has been devoted to research on the role and use of images in scientific practice (Lynch & Woolgar 1990; Beaulieu 2000; 2002; Alač, 2004). At the same time, sound has been relegated, if not largely forgotten (Cartwright, in preparation).

In an fMRI scanning, the attainment of experimental data is not directly dependent on sound per se. The high decibel “acoustic noise” is due to the vibration produced by the interaction between the static magnetic field and the time-dependent currents in gradient wires. Such noises, as already mentioned, may cause severe impediment for successful research. However, the following excerpt from an fMRI apprenticeship shows how reenactment of sound can play an important role in construction of *professional vision* (Goodwin, 1994).

The interaction described takes place between a new member of the laboratory—Nick (N), and the experienced practitioner Eric (E). The new member is being taught the procedure of fMRI data analysis. As is the usual practice (observed in other laboratories as well), Nick’s first experience of the fMRI procedure consists of his playing the role of an experimental subject. Before the analysis of the experimental data, he was scanned in the fMRI facility. In this process, lab members have an experimental subject, while Nick is more engaged in the viewing and analyzing of data, furthering his understanding of the procedure.

The excerpt below details a part of the interaction where the two practitioners are seated in front of a laboratory computer, and Eric explains to Nick how the data are organized:

EXCERPT 1

E: Ok there are three studies (0.2) tha:t (0.1) were in this
directory.//There was,
N: //Three?
E: Yeah. (There) = cause there is localizer. Do you remember at the
very //beginning
N: //At the beginning of functionals?
E: No at the beginning of the //MP-rages
N: //Oh ok
E: You remember you've heard sort of clicking ch-ch-ch ((*imitates the*
sound)) and then you got just a quick ah ba:um, ba:um, ba:um,
((*imitates the sound*)) and then it was quiet and G. said this
was the warm up and here is the real one? That's the localizer. **[24]**

In the beginning of the scanning process, the experimental subject is placed in the MRI scanner. After she or he is *aligned* so that the head is at the center of the magnet, a *localizer scan* is collected. The scan gives the exact location of the brain. By querying the

---

[24] 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.

number of the studies in the directory (line 3), Nick demonstrates uncertainty about the meaning of the localizer scan.

Eric, rather than providing a detailed linguistic and technical explanation of what a localizer scan is and when it is taken, imitates the sound produced by the scanner (lines 9 through 10). His voice, by reenacting the *acoustic noise* that accompanied the scanning process, arouses a recall of the experience. While Eric involves his body to provide an explanation of the professional category—the localizer scan—Nick's comprehension is achieved through an embodied understanding. In this process, he re-enacts the experience of being in the scanner. Eric's vocalizing works as a *coding schema* that organizes the messy world into categories relevant for the work of the profession (ibid.). Moreover, while it constructs objects of knowledge and rationality, it is neither linguistic nor graphical in form; rather, it is a nonlinguistic, acoustic event.

In this manner, what was an undesired aspect of the scanning session is reintroduced into the process of data analysis to assume an explanatory function. The confounding factor becomes an important element in the embodied procedure of sense-making.

4. CORRECTING FOR MOTION BY ENACTING MOTION

In addition to sound, another confounding factor—motion, and its role in understanding—is analyzed in the text that follows. Similarly to the previous section, this portion of the text reports on an apprenticeship interaction. Once again, the experimental data were previously collected, and the apprenticeship takes place through the process of data analysis. However, the activity described here is situated in a cognitive neuroscience laboratory that largely practices scanning of a brain damaged patient population. Hence, the experimental subject of the scanning session was not the practitioner herself, but a clinical patient. The interaction reports on practitioners' involvement in assessing the existence of artifacts in the experimental data.

Over the course of the experiment a series of brain images are recorded. Each scan or image represents a brain slice. During the preparation of the experimental data for statistical examination, among other types of manipulation, researchers need to asses if the brain slice representations can be aligned. The assessment is achieved by viewing slices (in the axial, sagittal, and coronal view) represented on the computer screen over the time-course of the experiment (Figure 3-2). The researchers, by using mouse commands, alternate the view of individual images on the screen. The goal is to check if the representations of the brain slices align with each other. When nonalignment is detected, the researchers try to correct it. This is done in order to prevent the appearance of visible defects in the images (i.e., blurry splotches) once the statistical analysis has been performed. When the nonalignment is pervasive throughout the data, the data set has to be rejected.

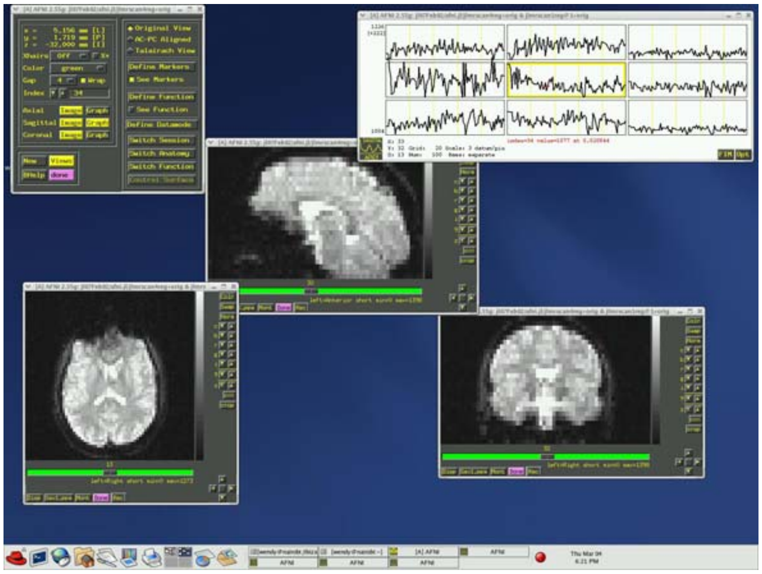


FIGURE 3-2: The computer screen displays brain slices over the time-course of the experiment

As the excerpt from interaction will attest, the nonalignment between the brain images is explained in terms of the subject's movement that caused it. The artifact, caused by subject’s movement, is called the *motion artifact.* In motion artifact the movement of the subject is considered to be an *artifactual cause*, rather than the *visible regularity of independent “natural” phenomena* (Lynch, 1989). While the subject’s brain and its function (evoked by experimental stimuli) are seen as independent natural phenomena, the potentially unintentional movement of the body is considered to be the cause of an intrusion or distortion in the visibility of fMRI images.

Notice, however, that what is believed to be the cause of the nonalignment (i.e., the movement) can only be inferred from its effect (i.e., the nonalignment of images). Rather than simply anchoring this nonalignment in the movement that caused it, the subject’s movement is inferred from nonalignment. While the images are present on the computer screen and can be directly compared, the movement of the subject could not have been seen or experienced by the researchers while the subject was laying in the scanner. The movements that the subject produces during the scanning section are usually too small to be seen on the computer monitor where MRI technicians and researchers supervise what goes on in the scanner. While in the previous section the apprentice’s prior experience was potentially evoked, in this example the experience is entirely created during the data analysis activity. Since the subject of the scanning session was a clinical patient and not the apprentice herself, the researchers create, rather than re-create the experience of being in the scanner.

Therefore, when considered from a purely theoretical perspective, what the researchers are dealing with is a circular explanatory flow between brain images and the subject’s movement. The images are not comprehended alone; they are understood in respect to the subject’s movements. Simultaneously, the researchers imagine such movements by grounding them in the brain images. Yet, are the images all there is for the practitioners to use in constructing their understanding of the nonalignment in terms of the subject movement? In other words, can the hypothetical movement be instantiated through some additional external structures and processes, other than brain images?

4.1. LEARNING TO SEE THE SUBJECT'S MOVEMENT

Excerpt 2 reports on twenty seconds of interaction between two fMRI practitioners. The practitioners are seated in front of a computer screen in the cognitive neuroscience laboratory (Figure 3-3). The room consists of rows of desks with computers, and is primarily used by the laboratory members for data analysis. The practitioners, by positioning and orienting their bodies toward each other and toward the computer screen, and by disregarding other events that simultaneously take place in the room (other lab members entering and exiting the room, and my videotaping them), determine the portion of the space that is relevant for their action. The person on the left is an advanced graduate student, Gina (G). Gina teaches Ursula (U)—a newcomer to the laboratory, seated on the right—how to analyze fMRI experimental data. The excerpt from the interaction was recorded during an early stage in the practice of data analysis where experimental data need to be prepared for statistical examination. Each image on the computer screen represents a brain slice at a particular moment of the experiment. Ursula's task is to check for alignment between images in the series (representing the time-course of the experiment). This is done in order to keep experimental data clean from motion artifacts.

EXCERPT 2

1 G: That's definite I can see her in this plane=

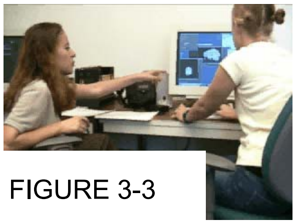


FIGURE 3-3

((*Points to the computer screen*))

2 going from here to here.


FIGURE 3-4


FIGURE 3-5

((Sweeps her arm in a downward motion and halts at a certain point))

3 Aaaa ((disapproves))

4 U: (Aaaa; this is good one) *((Points to the sagittal view of the brain slice on the screen))*

5 G: Slice, edit *((instruction for pressing buttons while manipulating computer screen))*

6 U: (Oh, here we are)

7 G: Is she moving any more in thirty? *((Moves closer to the screen))*

8 U: She's moving - =


FIGURE 3-6

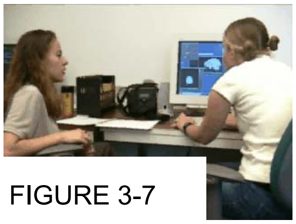


FIGURE 3-7

((Hunches))

9 Like this one


FIGURE 3-8

((Points to the axial view of the brain slice on the screen))

10 Is going down. Hhhhh ((laughs))


FIGURE 3-9

((Swings with right hand (palm down) downward))

What the excerpt illustrates is a *distributed cognitive system* (Hutchins, 1995) at work. The task at hand is being dealt with through an action distributed across the material and social world of practice. Gina and Ursula are not just looking at the computer screen to identify the non-alignment among brain slice representations. Instead, the apprenticeship proceeds as a part of an actual action. The process of cleaning the series from inadequate images is done through mouse commands and manipulation of the computer screen. Furthermore, the action is distributed across social actors – Gina and Ursula. Importantly, Ursula, the junior practitioner, controls the computer screen. She holds her hand on the mouse, and manipulates the computer display. However, it would not be entirely adequate to say that Ursula performs the action. Ursula's action is closely related to Gina's knowledge of the procedure. Left to her own devices, Ursula may not be able to see the difference between brain images, and hence perform the deletion of non-aligning ones. In addition to describing the distribution of the cognitive task across the material environment, the analysis that follows will focus on its distribution across multimodal semiotic resources. At the same time, for the sake of clarity, the analysis will treat the production of semiotic action as generated separately by each practitioner.

First, I will look at how the distortion of fMRI images is performed through Gina's gestural enactment. Gina's gesture produces visibility in the three-dimensional space of social action, and allows Ursula to learn about digital images by observing the embodied action. Next, I will describe how Ursula herself produces embodied interpretations of the images. She does so to comprehend image distortion through her bodily engagement with the digital screen. Finally, I will point out that the embodied interpretation of the images does not function *on its own*, or in *parallel* with the brain images. Instead, brain images and their interpretations build new, compound forms. These forms, while encoding an indeterminate referent, and appearing to have a semantically illogical character, play crucial roles in generating understanding for the young practitioner.

4.1.1. FROM VISUAL REPRESENTATIONS TO EMBODIED ACTION

Throughout the excerpt, Ursula uses mouse commands to manipulate brain images on the computer screen. The action of alternating the image display creates *apparent motion*, or the impression that the distorted image is moving. The existence of apparent motion is reflected in Gina's use of motion verbs: "… going from here to here" (lines 1-2). However, Gina does not only speak about the movement of the images. Rather than simply seeing the difference between the two images, Gina "sees" the experimental subject moving: "I see her in this plane". By changing the display of a particular brain slice through time, the movement of the subject becomes *visible*. If the subject's body and behavior were yielding themselves to the prescribed procedure, *she* (the experimental subject) would be promptly cancelled or translated into abstraction. But because of the subject's disruptive movement, *she* becomes visible. Yet, her visibility is recognized during the experimental data analysis procedure just so that what are believed to be the consequences of her undesired behavior (i.e., the motion artifact) can be removed from the analysis.

But, is the harmful but explanatory movement only understood from comparing brain images on the computer screen? To address this issue, we need to carefully look at the unfolding of discursive action, as it reveals a pervasive use of gestural performance. As shown in line 2, Gina quickly points toward the computer screen (Figure 3-3), and consequently morphs that deictic sign into the gesture of the subject's movement. To achieve this end, Gina's left arm, palm downward, outstretches to a position in front of her. It sweeps downward and halts at a certain point (Figures 3-4 – 3-5).

The initial deictic gesture, coordinated with Gina's and Ursula's body orientation, determines the center of attention – a particular brain image on the computer screen. Nevertheless, Ursula may not yet *see* the inadequacy of the identified digital image. The inadequacy has to be reproduced in an additional – hopefully, more immediate – form. Even though the practitioners could resort to other, previously existing, stable visual representations (such as brain images reproduced in a brain atlas), they opt for more ephemeral, but concrete, online generated structures, based in action. The digital brain representations have to be coordinated with such future structures. The initial deictic gesture establishes the linkage between the two. By executing the deictic gesture, and consequently, by mapping the subject's movement onto the gesturing hand, the gesture *liberates* the apparent movement from the computer screen and enacts it in the three-dimensional space. The subject's movement *migrates* from the realm of visual representations into the world of embodied action.

The gesturing performance does not simply replicate what is already present on the computer screen. The performance is an interpretation of the digital representation – it functions as a solution to the problem that the representation presents. The horizontal, rather than vertical position of the gesturing hand shows that the gesture doesn't merely stand for the brain images present on the computer screen. Instead, Gina's hand represents the moving head of the subject[25]. In other words, the gesture is not just about the movement of the images; it is about the subject moving.

However, the idea of subject moving is already present in Gina's linguistic expression ("I can see her in this plane going from here to here", lines 1-2). Why is then Gina gesturing? Is her gesture purely redundant?

To confront this dilemma one could imagine a situation where the gesture was not executed, perhaps if Gina just said that she saw the subject moving while looking with Ursula at the apparent motion of the digital images. What would such a situation require from Ursula? Ursula may be confused by Gina's claim that she *sees the subject moving*, when what she is really seeing are *brain images moving*. In order to see the subject moving while looking at the apparent motion of brain images, Ursula would need to compute a complex interpretative leap. She would need to mentally construct a three dimensional movement of the subject lying in the scanner from two-dimensional images of brain slices. Gina's talk, even though it indicates such movement, doesn't reproduce it.

The gap between the brain images and Gina’s seeing subject’s movement in such images is bridged by her gesture of the subject’s movement. In contrast to the apparent movement of the two-dimensional images, the gesturing hand and the subject’s head movement share a larger number of characteristics, including the hand’s production of a three-dimensional horizontal movement, similar to the supine body. Because of the mobility of the hand, the mapping between the gesture and the subject’s movement is more immediate, which reduces complexity of Ursula’s interpretative task. Moreover, the performance of the subject’s movement generates concreteness and enables Ursula to observe in the three-dimensional world of action what Gina claims to see on the computer screen. Rather than understanding the existence of the artifact solely through a comparison and alignment of two-dimensional images, Ursula learns about fMRI brain representations, by observing embodied performance of the subject’s movement unfolding in the shared world of action.

Furthermore, through the multimodal action, the individual and particular reappears. In contrast to the impersonal style of scientific papers, as well as the use of fMRI data for general claims about human brain and cognition, at the stage of data analysis described here, Gina and Ursula employ the brain representations to discuss, and enact the behavior of a single experimental subject. The recreation of particular movement of the subject is instantiated in the public space of action, rendering the abstract data comprehensible.

4.1.2. EXPERIENCING THE SUBJECT’S MOVEMENT

The analysis developed in the previous section focused on Gina’s embodied involvement in the meaning production. Yet, the focus on Gina’s activity doesn’t imply Ursula’s passivity. Ursula doesn’t just acts as a conduit through which Gina extends her action on the computer screen. By observing Gina’s gesture and by manipulating computer screen, Ursula

---

[25] This claim is corroborated by the representations of the subject’s movement simultaneously produced in the linguistic form, as well as through body movements, discussed in what follows.

actively participates in the embodied process of learning. Her active involvement in such a process is further elaborated in this section.

Gina’s gesture in line 2 is a clear example the subject’s movement enactment. Yet, such an enactment becomes still more elaborated through Ursula’s action in lines 7-10. By involving her hand along with her body movement in the semiotic process, Ursula applies additional layers of meaning onto the digital brain images. Ursula’s active use of multiple embodied modalities corroborates the claim that such a use is not only peculiar to the role of teacher, nor it is accomplished for the sake of demonstration only. Instead, it is also employed to reason about the examined phenomenon. Ursula’s performance illustrates reasoning as a social process grounded in the body.

The unit of action (lines 7-10) opens on Gina’s initiative. In line 7, she moves closer to the computer screen (Figure 3-6) and asks: “is she moving any more in thirty?” Gina’s movement signals that her question is not rhetorical. At the same time, her increased proximity to the computer screen functions as an attention channeling device that directs Ursula’s gaze toward the screen. The body movement assumes a role similar to the deictic gesture performed in line 1, Figure 3-3. In response to the question, Ursula takes the floor and actively engages into creation of subject’s movement (lines 8-10).

To signify subject’s movement, Ursula combines three types of semiotic forms with brain images and talk (lines 8-10). In line 8, while her hands are engaged with the keyboard, she hunches (Figure 3-7). In line 9, she points toward the computer screen (Figure 3-8). Finally, in line 10, she executes a hand gesture where she swings with right hand (palm down) downward (Figure 3-9).

The body movement performed in line 9 is of particular interest. While hunching, Ursula’s body assumes a crucial role in the meaning-making process. Similar to Gina’s gesture in line 2, Ursula’s movement becomes a three-dimensional enacted sign of the subject’s behavior in the scanner. The interpretation involves not only a part of the body, but the body as a whole.

Once again, Ursula learns about the non-alignment of fMRI images by observing the enactment of its hypothesized cause. Rather than understanding the non-alignment solely through a comparison of visual features on the computer screen, the body movement enables Ursula to understand and learn to see the non-alignment by observing the embodied performance of its cause.

Ursula's body movement doesn't function in isolation. Instead, the meaning of the discursive act is produced through mutual coordination among various semiotic modalities and the activity in progress. To identify acts that produce meaning, and their developments in the world of action, the attention should be centered on the interaction among Ursula's body movement, her talk, fMRI brain images, and her manipulation of the computer screen.

Ursula's linguistic expressions confirm the interpretation of the body movement in terms of the subject's behavior. The expression "She is moving" parallels the hunching movement and expresses the idea of the subject's body in motion. The apparent movement of the brain images and the assertion that the subject is moving are bridged with the direct involvement of the body. Ursula's body movement enables practitioners to jointly see the subject's movement in the environment of practice. At the same time, the production of meaning through the body movement is directly dependent on talk; through talk, the body movement becomes comprehensible. The linguistic expression makes apparent that Ursula's body in line 8 isn't presented only as her own body that hunches. Instead, it also stands for the specific movement that the subject produced in the scanner.

The semiotic function of Ursula's body is directly shaped by the activity in progress. While her engagement with the keyboard doesn't permit a gesturing performance, such an activity is compatible with the hunching movement. The type of semiotic performance employed by Ursula in line 8 makes a production of meaning at multiple channels possible: Ursula can execute the body movement in parallel with the apparent motion of images.

This, in turn, allows for a mutual coordination among the apparent motion (produced on the computer screen) and the subject's movement (executed in the environment of practice). The image manipulation guides the hunching movement step by step while the hunching provides visibility and concreteness for what is detected in the images. Ursula's body and the digital display shape each other reciprocally. The body learns to see what is in brain images (Latour, 2004), while the brain images become visible through their enactment.

Even though the body motion shares additional characteristics with the subject's movement, such representation is further removed from what the scientists have to act on: the task is to select and computationally correct the brain images. Therefore, the representation of the moving body has to be linked back to the images. This is achieved through two hand gestures: a deictic gesture (Figure 3-8), and a gesture that enacts the subject's movement (Figure 3-9).

After executing the body movement (line 8), Ursula identifies the non-aligned image on the computer screen (line 9). Once the image is identified, she removes her hand from the keyboard, and performs a gesture of the subject's movement (line 10).

In line 9, the deictic gesture combines the subject's movement (enacted through hunching, line 8) with the digital brain image. Once again, the deictic gesture, similar to Gina's performance in lines 1 and 7, links the three-dimensional action with the screen; it assumes the role of a mapping process between the moving body and the brain images.

In line 10, the deictic gesture is swiftly transformed into another hand gesture of the subject's movement where Ursula's arm maps onto the subject's head. Even though the subject's movement was already performed through hunching, the hand gesture is not redundant. While the hunching has the advantage of preserving the straightforward mapping between elements of two domains (Ursula's body parts map onto the subject's body parts: shoulders map onto shoulders, head maps onto head), the hand gesture can signify the direction of the subject's movement more precisely. Because of its orientation and shape, the hand gesture contains some deictic elements that the hunching movement doesn't: the gesture directs practitioners' attention back to the computer screen.

Similarly to the activity analyzed in the previous section, the description of the action performed in lines 7 through 10 illustrates how participants generate understanding of fMRI images through their performance of the subject's movement. In the previous section such movement was performed by the senior practitioner. Here, Ursula—the newcomer in the laboratory—performs the movement. In addition to staging the subject's movement through her gesture, Ursula enacts the movement through hunching.

The hunching assumes various functions. Like the gesture, it translates the apparent motion from the brain images into the movement of a human body, and it makes Ursula's interpretation of the experimental data publicly available. When Ursula hunches and aligns her own body with the digital images, she makes the disturbance of the fMRI data visible to Gina. By physically enacting the process of change seen in the brain images, Ursula employs her body to create the hypothesized movement. The movement is necessary for an adequate explanation of the disturbance attributed to the experimental data. By making the images visible through her embodied performance, Ursula signals her understanding. She doesn't only inform Gina about the artifact – she indicates her mastery of the task.

Furthermore, the analysis of the interaction between Gina and Ursula shows how Gina acquires professional knowledge not only by looking at brain images, or by following formally specified procedural rules. Instead, she learns through an active involvement of the body.The gesture, as well as the body movement, participate in closing the interpretative gap between the non-alignment among images and their causes. Gina doesn't have to only mentally interpret the non-alignment in terms of the subject's movement. Her observing of the subject's movement through its enactment is a crucial component of such an act. The gesture allows for the cognitive process to be distributed across the environment of practice (Hutchins, 1995). The process is distributed not only across stable representations (such as digital brain images), but also across ephemeral forms generated through the practitioner's body.

The description of the hunching movement allows for an additional way of learning. It produces first-person understanding for its performer. Hunching allows Ursula to experience a movement similar to the one the experimental subject may have experienced in the scanner. While the gesture creates the idea of the subject moving that is directed towards the images and functions as an event that can be observed by the practitioners, the hunching generates a model of the subject's movement that can be experienced in the first person. In this sense, Ursula's body movement is used to communicate, as well as to make sense of the problem to be solved. It illustrates how the newcomer learns to read brain images through the experience of her own body in motion.

Moreover, the analysis underlined a further point of interest. It indicated the importance of the coordination between modalities. Such coordination is crucial for the production of meaning. To understand how embodied modalities work the analysis has to be directed at their coordination with talk, brain images, and the on-going activity. It was pointed out how body movements in line 7 and deictic gestures in line 9 instantiate mapping between images and the subject's movement. The uncovering of coordination among modalities enables further insights into meaning-making processes realized through practice. Such processes will be the focus of the following section.

4.2 WHERE IS THE ARTIFACT?

The description of the activity reported in the previous two sections illustrates how learning to identify a nonaligned brain slice representation takes place through an embodied recreation of the imaginary subject's movement (lines 1-2 and 7-10). Ursula is guided in spotting the nonaligned representations by looking at and reenacting movements. The reenactments represent those movements that practitioners believe the experimental subject made during the scanning session. Ursula is asked to understand the discrepancy between images in terms of the hypothesized subject's movement that caused it.

This evocation of the movement can be detected in the on-going talk. Gina verbally explains the nonalignment of the images by using linguistic forms that express motion. In lines

2 and 10 the verb “to go” is used, while in lines 7 and 8, Gina talks about *her moving*. These expressions of movement are in clear contrast with the static character of the brain images.

However, a closer look at Gina’s talk reveals something unexpected. Gina’s utterance, while evoking the subject’s movement also expresses spatial location. This is noticeable in lines 1, 2, and 7: “I can see her in this plane” (line 1); “going from here to here” (line 2); “Is she moving any more in thirty?” (line 7). Note that the expressions “from” and “here” do not stand for the physical location in which the subject’s movement effectively took place; rather, such expressions indicate the brain image[26]. This is even more clearly visible in lines 1 and 7, wherein Gina talks about “this plane” and “in thirty”. The expression “from here to here”, the idea of “the plane”, and the number thirty all denote brain images, rather than the subject who is “moving”, “going”, and “being seen”. At the same time, they are directly combined with the linguistic expression “she”, referencing the subject who had moved physically in the scanner[27]. The subject becomes *visible* or *moves in a particular brain slice representation*. Yet, such a construct is illogical[28]: the experimental subject cannot move in a static representation of her brain slice”.

The usage of expressions “in this plane”, ”from here to here”, and “in thirty” in lines 1, 2, and 7, supported by the general ethnography of the field, indicates that the combination of the realm of the subject’s movement with the idea of the brain images not only announces the linkage between the two (Gina does not only speak about the subject *and* the images), but reflects a construction of a *hybrid semantic form*. This form simultaneously represents the subject’s physical movement and the nonalignment of the brain images. By dealing with something that is concurrently about what exists in the experimental data and stands for the

---

[26] This interpretation is backed up by 1) the knowledge that the temporal sequence of images is represented on the computer screen in a spatial array (there is the time-spatial sequence mapping) so that “when” can be expressed by “where”; 2) the subsequent analysis of participants’ orientation toward the computer screen and their use of indexical gestures toward the brain images.

[27] This interpretation will be corroborated by the analysis of participants’ iconic gestures that represent the subject’s movement in the scanner.

[28] For and identification and discussion of a similar phenomena see Ochs et al., 1996.

concrete movement of the subject, the researchers create and manipulate a hybrid and rationally impossible entity[29].

The construction of such an entity is crucial for the practice: it allows the practitioners to understand the motion artifact*.* The analysis of the practitioners' talk illustrates how the artifact is not the subject's movement per se; it is the movement as reflected in the images. The disturbance is not present in any single image, but it exists in the difference between the images. This difference is understood in terms of the subject's movement. The artifact stands at the same time for the subject's undesirable movement explaining the images appearance, as well as the brain images where such movement is spotted. Once Ursula spots the *subject's movement in the images*, she can identify the movement artifact. In other words, the process of hybridization is vital for understanding of the experimental data.

But is the production and existence of such a hybrid structure confined to the inner workings of the individual minds?

The previous two sections illustrated how not only brain images, but also how the imaginary movement of the subject can be largely constructed in the environment: practitioners use their bodies to perform the subject's movement. The same is true for the process of coordination between brain images and the subject's movement, such a process can be instantiated in the environment of practice as well. Deictic gestures are not only

---

[29] The existence and mechanism of such hybrid conceptual entities expressed in linguistic forms was first explained by *Conceptual Integration Theory* (Fauconnier & Turner, 1998, 2002). Similar to the well known *Conceptual Metaphor Theory* (Lakoff, 1987), Conceptual Integration Theory unifies the analysis of metaphor with the analysis of other linguistic and conceptual structures. This theory is involved in identifying systematic projections of language, imagery, and inferential form. However, its peculiarity lies in its focus on novel and often short-lived conceptualizations. Instead of accounting for stable knowledge structures represented in long-term memory, it models the dynamic evolution of speakers' on-line representations (Coulson & Oakley, 2000). According to Fauconnier and Turner, conceptual integration or blending operates on mental spaces. Mental spaces are short-term mental scenarios constructed on-line. In blending, elements from input spaces are projected to a separate space, the "blend". The blend inherits partial structure from the input spaces, and has emergent properties of its own. Blending is an online, real-time process that creates new meaning through the juxtaposition of familiar material.

devices for indicating particular features in the environment, but actions that allow for mapping and integration between domains. In lines 8-9, Gina first enacts subject's movement through her hunching. Consequently, she points toward the computer screen. This deictic gesture does not just select an image on the computer screen. Through such an action the embodied instantiation of the subject's movement is coordinated with the brain image. The deictic gesture functions as a mapping process between the two domains.

In addition to deictic gestures, mapping can be also grounded in an active use and manipulation of space and time. In line 2, Gina coordinates her talk ("going from here to here") with a gestural form representing subject's movement. An examination of her gesture and its mode of execution in the context of the practice reveal that the gesture itself does not express a unitary form. The marked halting of the gesture (Figure 3-3) indicates the imaginary location to which the subject's body moved while in the scanner. Importantly, however, the same component of the gesture stands for one of the unaligned brain slice representations. The gesture, through its timing (continuous movement versus halting), connects the subject's motion with the brain image to produce the hybrid construct. The coordination among the gesture, images, and the gestural timing generates the idea of the *motion in the images*.

Also the use of space can be recruited into the mapping process. In line 10, after performing the subject's movement through her body and pointing toward the computer screen, Ursula enacts the subject's movement through her hand gesture. By transferring the representation of subject's movement from her body onto her gesturing hand, she physically moves such representation closer to the computer screen. The spatial proximity between the two realms allows for their mutual contamination and the production of the motion artifact. The motion artifact is a veridically impossible structure assembled through the publicly shared action.

It has been documented that accounts of artifacts play a crucial role in studying science as a practice (Lynch, 1985). They function as "…detailed showing of a work environment as part of specific empirical displays of data" (ibid.: 85). According to Michael

Lynch's distinction between *positive* and *negative* artifacts, the motion artifact here described is a positive artifact. Lynch describes positive artifacts in the following manner:

> 1. They have a thing-like substantiality, visibility, and availability to accounts. As such they are described as particulars in the revealed … scene which interfere with the transitivity of the practical access to "natural structures".
> 2. They are "common" and regular in their appearance as the everyday troubles that inhabit the work of doing technical preparations of displays.
> 3. Once they are located as artifacts they are unproblematically separated from the adequately constructed features of a display format …
> 4. The discovery of these artifacts initiates accounts directed not always to their "source" in the practical work, but more importantly to the immediate issue of whether the material should be used in light of their accountable presence. (ibid.: 91).

Though this definition suggests that the motion artifacts may be characterized as *readily available to instructed looking, while they appear as "things" in an observed field* (ibid.: 91), the discursive action of two fMRI practitioners urges us to open up such portrayal to further investigation. The idea of motion artifact's *availability to looking*, and the belief in its *thing-like appearance* are products of an assumption that visual representations are directly legible through their detached viewing.

Two previous sections illustrated how the availability to looking is a result of coordination among various semiotic modalities. The practitioners' accounts of the artifact reveal an understanding of the nonalignment through enactment of what is believed to be its cause: a subject's movement. The visibility is realized not only through looking, but through an embodied performance of the subject's movement.

Furthermore, the analysis of the coordination between the talk and embodied interaction reveals that practitioners do not understand artifacts as entitles that reside in particular locations, be it brain images, or gestural and body enactments, from where they can be disclosed. On the contrary, the artifact isn't a referentially simple, well-defined object that waits to be spotted. Instead, the artifact dwells between various semantic domains, and emerges through coordination of clashing elements (e.g., it is at once two and three-dimensional, static and dynamic). Moreover, the attention to multimodal interactions shows how such coordination isn't exclusively an internal product of a single individual, but rather a distributed process generated through an active, bodily involvement. The artifact is an emergent property generated through such an involvement, and performed in social and physical environment of practice.

5. CONCLUSION

In this paper I have portrayed science as a phenomenon entrenched in specific everyday instances of its practice (e.g., Lynch, 1985; Latour, & Wollgar, 1979; Knorr-Cetina, 1981). The scientific practice described is fMRI brain mapping. I focused on instances of the practice by following leakages. The leakages allow for the unwanted, and usually forgotten, elements of fMRI procedure to pour back into the practice. The unwanted elements are acoustic noise of the scanner and movements of the experimental subject. While fMRI practitioners judge such elements to be extremely harmful for the scientific procedure, they emerge here as beneficial. They play an important role in apprenticeship activities where the re-production of sound and movement generates visibility and understanding.

The tracing of how unwanted elements pour back in the fMRI practice implies an abandonment of a language-centered analysis (or any analysis focused on just one semiotic modality, such as gesture, e.g. McNeill, 1995). Instead, the analysis was focused on how practitioners interact with technology while using and coordinating multiple semiotic modalities (Goodwin, 1994, 1995, 2000, 2003).

I began this paper by asking how interpretation of fMRI brain images takes place. I wondered if the understanding of fMRI images has to be entirely accomplished through their detached looking. The description of two instances of multimodal interaction reveals that the interpretation of digital brain images is rather embodied, and that seeing is constructed through action. To learn how to spot inadequacies in experimental data, the young practitioners coordinate digital brain images with talk, as well as gesture, gaze, non-linguistic sound, body movement and orientation. Insofar as practitioners, in order to interpret images and actively engage with them, one cannot talk about a *visual autonomy* (Hansen, 2004: 26) in fMRI. The meaning of brain images does not lie in the images themselves, nor is it entirely lodged in the visual system of fMRI practitioners. Instead, seeing is a process *contaminated* by the semiotic body in action.

Related claim regards the assumption that *the reality fMRI images reveal is impersonal.* The pervasive reappearance of the experimental subject in the process of data analysis indicates that practitioners work with fMRI images by enacting what is personal and individual. To understand brain images, they evoke the body of the experimental subject. The evocation is accomplished via engagement of practitioners' bodies.

In fact, I wrote about semiotic bodies throughout the text. I was concerned with the appearance of the body of the experimental subject and the bodies of the practitioners. By engaging their bodies, the practitioners listen to the recreation of the scanners' noise, and enact the hypothesized movement of the subject's body. The enactment of the scanner noise allows for re-creation of the experience of being in the scanner, while the use of body movements generates understanding through creation of a hypothetical experience (of being the experimental subject). The emergence of bodies through the contextual practice of science is curious. It contrasts the fate that such bodies have in the final accounts of scientific work. Once the results of a scientific practice move outside the laboratory (in form of scientific articles, conference talks, or newspaper reports), those bodies become entirely transparent[30]. Yet, the two excerpts illustrated how subject's body, as well as the bodies of the practitioners, can be vital in the actual laboratory practice. Such bodies function as critical factors in the processes of meaning production.

How can the pervasive appearance of the subject's movement in the data analysis be explained? According to one possible explanation, the subject's movement is a product of a

---

[30] In fact, if I were to analyze different types of documents, such as verbal accounts or published scientific reports, for example, rather than analyzing audio and video-taped records of scientific practice, the embodied performance would largely remain invisible.

conceptual conversion, where scientific representations are translated to concrete reality. Proponents of cognitive semantics have noted that humans tend to ground understanding of complex phenomena in embodied experiences (Lakoff, 1987, Fauconnier and Turner 2002). Because of its sophistication, understanding of technical phenomena is often based upon human biological capacities and experience of functioning in a physical and social environment. In motion correction practice, by enacting movement, practitioners convert experimental data in concrete events: fMRI images are understood in terms of a person's body moving in the scanner. While young practitioners need to learn how to identify differences between fMRI images, it is quite immediate to understand a movement of the person in a lying position. Through their translation into the subject's movement, the experimental data acquire an unequivocal, real-world referent.

The idea of conversion into the concrete, unequivocal world referent is definitely not misleading; still, the analysis of situated practice reveals a more complex reality. Practitioners cannot just translate the inadequacies of the images into the causes of such inadequacies. They need to look at images while seeing the subject's movement. How is this accomplished?

The attention toward multimodal, situated practice indicates that what practitioners deal with is not simply a translation of brain images in subject movements. Instead, we see a process of contamination between bodies and images. In this respect, the motion artifact is not a thing or a well-defined location. Rather, it reveals itself in the linkage between different semiotic modalities and the material environment of social action. The appearance of such unusual, imaginary constructs in scientific practice is not only reserved for *extraordinary acts of creation* (Fauconnier & Turner, ibid.). The imaginary constructs appear in routine sessions of laboratory apprenticeship. Ursula's acquisition of motion correction expertise proceeded through processes of contamination and integration between apparently irreconcilable meaning domains. Such complex processes of understanding and knowledge acquisition are largely generated and performed through social action.

In summation, the text points out how practitioners use their bodies to shape the images for their understanding. Additionally, the focus on multimodal practice reveals that, rather than being exclusively based in simple, referentially unambiguous concepts, scientific apprenticeship concerns processes of imagination, and contamination across meaning domains. The excerpts from interaction show how such domains can be instantiated and coordinated through public, multimodal re-enactments of what are considered to be unwanted elements in fMRI procedure.

ACKNOWLEDGEMENTS

I would like to thank the following people for their contribution to this project: Wendy Ark, Geoff Boynton, Lisa Cartwright, Aaron Cicourel, Seana Coulson, Gilles Fauconnier, Deborah Forster, Charles Goodwin, Marjorie Harness Goodwin, Ed Hubbard, Edwin Hutchins, Silvia Paparello, Marty Sereno, Jack Sidnell, Joan Stiles, and the participants in the ethnographic study.

WORK CITED

ALAC, M
2004 "Negotiating pictures of numbers", *Journal of Social Epistemology,* 18:2: 199-214.

ALAC, M. & HUTCHIN, E.
2005 "I see what you are saying: An ethnographic study of cognition as action in fMRI brain mapping practice", *Journal of Cognition and Culture*, 4:3: 629-661.

BEAULIEU, A.
2000 "The brain at the end of the rainbow: the promises of brain scans in the research field and in the media," In J. Marchessault & K. Sawchuk, (eds)., *Wild Science: Reading Feminism, Medicine, and the Media,* 39-52. London and New York: Routledge.

2002 "Images are not the (only) truth: Brain mapping, visual knowledge, and iconoclasm", *Science; Technology, & Human Values*, 27, 53-86.

CARTWRIGHT, L..
In preparation "An Industry Melodrama: Reading US Motion Picture Engineers' Writings on Film Sound Technology, 1930-1950".

COULSON, S. & OAKLEY, T.
2000 "Blending Basics" *Cognitive Linguistics,* 11-3/4: 1-14.

DUMIT, J.
2004 *Picturing Personhood: Brain Scans and Biomedical Identity, Princeton*, NJ: Princeton UP.

FAUCONNIER, G. & TURNER, M.
1998 "Conceptual Integration Networks", *Cognitive Science*, 22: 133-187.

2002 *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*, New York: Basic Books.

GOODWIN, C.
2003 "The Semiotic Body in its Environment", In J. Coupland & R. Gwyn *Discourses of the Body*, 19-42. New York: Palgrave/Macmillan.

2000 "Action and Embodiment Within Situated Human Interaction", *Journal of Pragmatics* 32: 1489-522.

1997 "The Blackness of Black: Color Categories as Situated Practice", In L. Resnick, R. Säljö, C. Pontecorvo and B. Burge (eds.), *Discourse, Tools and Reasoning: Essays on Situated Cognition*, 111-40. New York: Springer-Verlag.

1995 “Seeing in Depth”, *Social Studies of Science* 25: 237-74.

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

HEATH, C. & HINDMARHS, J
2000 “Configuring action in objects: From mutual space to media space”, *Mind, Culture and Activity*, 7, 1/2: 81-104.

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

HUTCHINS, E. & PALEN, L.
1997 "Constructiong Meaning From Space, Gesture, and Talk", In L. Resnick, R. Säljö, C. Pontecorvo and B. Burge (eds.), *Discourse, Tools and Reasoning: Essays on Situated Cognition*, 23-40. New York: Springer-Verlag.

HANSEN, M.B.N.
2004 *New Philosophy of New Media*, Cambridge, MA: MIT Press.

KNORR-CETINA, K.
1981 *The manufacture of knowledge: an essay on the constructivist and contextual nature of science*, Oxford: Pergamon Press.

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

LAKOFF, G.
1987 *Woman, Fire, and Dangerous Things*, Chicago: Chicago UP.

LATOUR, B.
1987 *Science in Action*, Cambridge, MA: Harvard UP.
2004 “How to talk about the body? The normative dimension of science studies”, in M. Berg & M. Akrich, *Bodies on Trial*, special issue of *Body and Society*, 10, 2/3, 205-229.

LATOUR, B. & WOOLGAR, S.
1979 *Laboratory Life. The social construction of scientific facts*, London and Beverly Hills: Sage.

LENOIR, T.
2004 “Foreword”, In M.B.N. Hansen.

LYNCH, M.
1985 *Art and Artifact in Laboratory Science: A Study of shop work and shop talk in a research laboratory,* London: Routledge & Kegan Paul.

LYNCH, M & S. WOOLGAR (eds.)
1990 *Representation in Scientific Practice*, Cambridge, MA: MIT Press.

McNEILL, D.
1995 *Hand and Mind: What Gesture reveal about Thought*. Chicago: University of Chicago Press.

OCHS, E., GONZALES, P., & JACOBY, S.
1996 “When I come down I’, in a domain state: Talk, gesture, and graphic representation in the interpretative activity of physicists”, In E. Ochs, E. Schegloff, and S. Thompson (eds.), *Interaction and Grammar*, 328 – 369. Cambridge: Cambridge UP.

ROEPSTORFF, A.
2004 “Mapping brain mappers: An Ethnographic Coda”, IN R.S. Frackowiak et al. (eds.), Human brain function (2nd ed.), 1105-17. Amsterdam: Elsevier.

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

SUCHMAN, L.
2000 “Embodied Practices of Engineering Work”, *Mind, Culture, and Activity*, 7 (1&2), 4-18.

STAR, S. L.
1989 *Regions of the Mind: Brain Research and the Quest for Scientific Certainty*, Stanford: Stanford UP.

This chapter is reprint of

Morana Alač: "From Trash to Treasure: Learning about the Brain Images through Multimodality", *Semiotica*, 156-1/4, 177-202, 2005.

CHAPTER 4

I SEE WHAT YOU ARE SAYING:
ACTION AS COGNITION IN FMRI BRAIN MAPPING PRACTICE

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.

1. 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 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.

2. 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 4-1). 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.

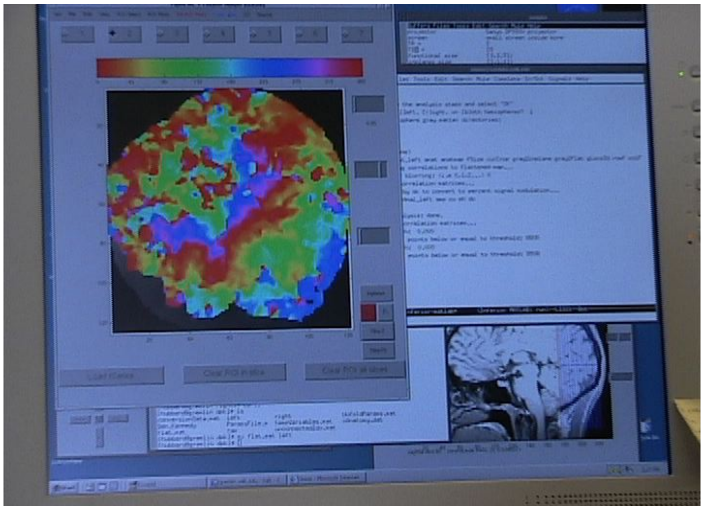


FIGURE 4-1: Phase map as viewed on the computer screen

3. 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.

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.

4. 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 retionotopic maps on the human cortex (i.e., to inscribe position and form of visual areas on a phase map, as shown in Figure 4-2). 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.

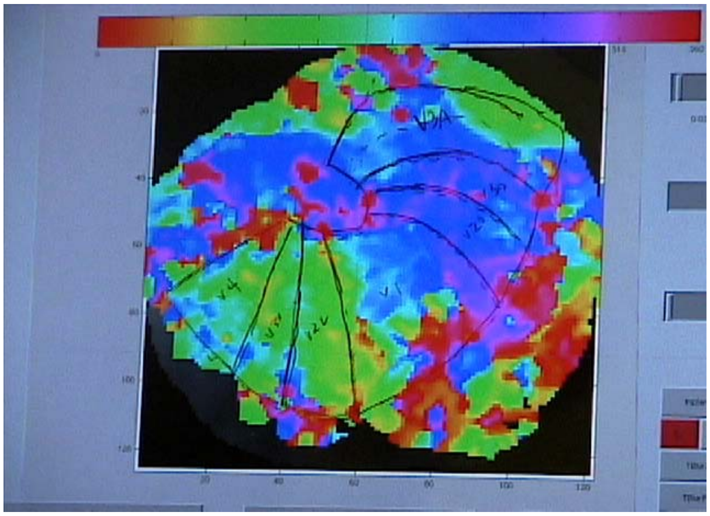


FIGURE 4-2: Phase map captured at a later stage of the practice after the expert and the novice had already inscribed on it borders of early visual areas

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 4-3). 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 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.

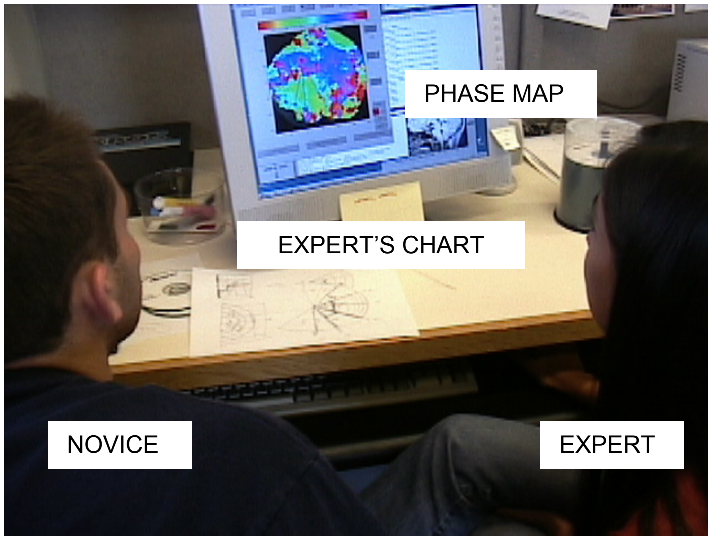


FIGURE 4-3: The setting

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.

5.1. SEEING THE EXPANDING RING

During the stages of "learning to see retinotopy", the actors view on the computer screen a phase map (Figure 4-1). They try to define early visual areas (i.e., V1, V2, V3) on that representation. The phase map is a static 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 4-4, 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.

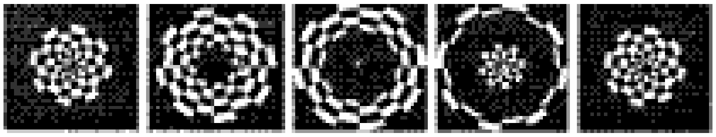


FIGURE 4-4: Expanding ring stimulus shown at five moments in time (from Engel et al., 1997)

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

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 <u>start</u> and actually see *((the novice moves closer to the screen))* the maps in
3 this noisy data[31]

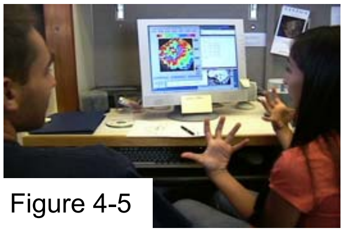


Figure 4-5

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

---

[31] 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.

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 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 4-5). 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.

5.1.2. SQUEEZING THE VISUAL FIELD

The meaning complex discussed in the previous section (and shown in Figure 4-5) suggests the idea that maps are somehow in the noisy data, and they can be seen by those who know hot 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). At the same time, the knowledge is also deposited in artefacts, such as the expert’s chart (Figure 4-6).

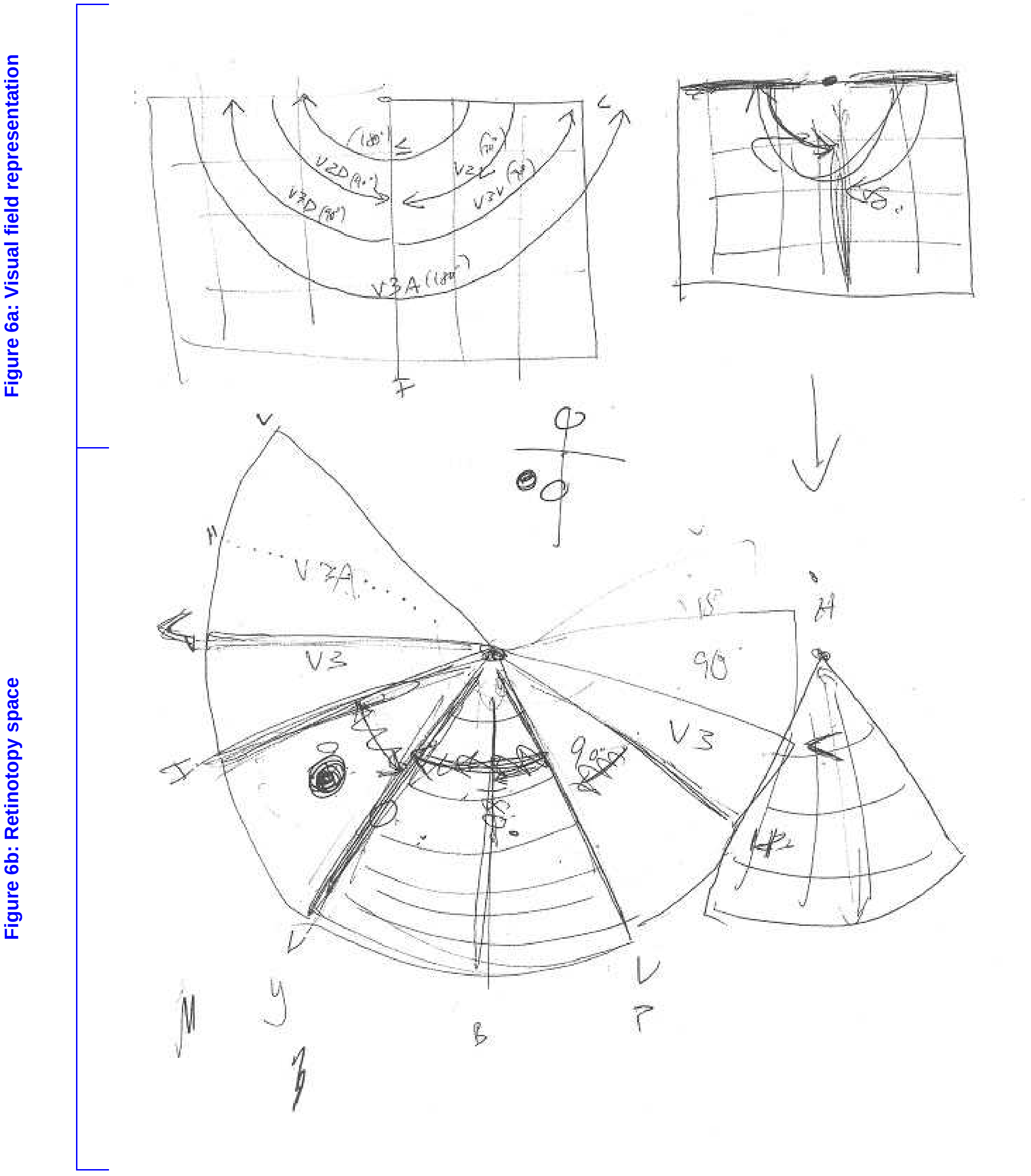


Figure 6a: Visual field representation

Figure 6b: Retinotopy space

FIGURE 4-6: The expert's chart

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 4-6). The chart can be divided into two parts: the diagram of the visual field

(Figure 4-6a), and the “retinotopy space” (Figure 4-6b). The retinotopy space represents the cortical organization of visual areas. By the end of the training session, 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.

Here is an excerpt from the interaction that took place during the chart construction. The expert has drawn the visual field (diagram on the right side of Figure 4-6a), but not yet its corresponding retinotopy space (diagram on the right side of Figure 4-6b):

EXCERPT 2

4 E: So take these two meridians and as just
5 as if you were squeezing them together into the pie shape

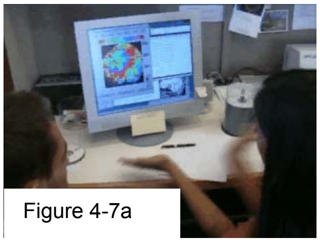


Figure 4-7a

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

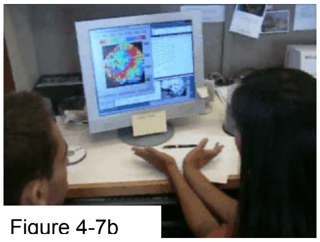


Figure 4-7b

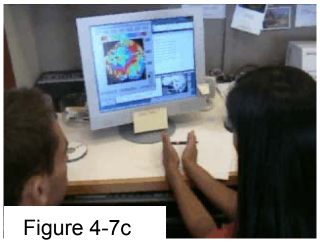


Figure 4-7c

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

((draws an arrow pointing to the right of the first drawing))
6 So you took this space
7 and turned it into that space ok? ((draws a gridded triangular shape))

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 it 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 4-7).

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.

5.1.3. 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) (Figure 4-6). 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.

5.1.4. 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:

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

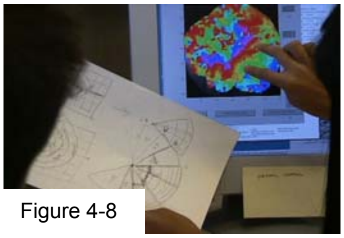


Figure 4-8

((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?

As the participants' discursive action described in section 5.1.1. (Figure 4-5) 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). Nonetheless, 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)[32]. Yet, while in the section 5.1.1. the mapping process was suggested through speach, 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 4-6)[33]. 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

---

[32] 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.

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

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.

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.

5.2. 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 4-9). Like the expanding ring, the rotating wedge stimulus ideally provokes the alternation frequency of the neuronal response that varies in the temporal phase.

FIGURE 4-9: Rotating wedge stimulus shown at five moments in time (from Engel et al., 1997)

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 traveling 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.

5.2.1. 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 4-2). 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*[34]) so that she can more convincingly justify the position of the borders on the brain map[35].

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 bacto 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

---

[34] 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.

[35] 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).

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 4-5). 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 section 5.1.2. we 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.

6. 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. In 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 sorts 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.

7. 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.

ACKNOWLEDGEMENT

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.

WORK CITED

BEST, M. & POCKLINGTON, R.

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

CLARK, A. & THORNTON, C.

1997 ‘Trading Spaces: Computation, Representation, and the Limits of Uninformed Learning’. *Behavioral and Brain Sciences* 20, 57-90.

ENGEL, S.A., GLOVER, G.H., WANDELL, B.A.

1997 ‘Retinotopic Organization in Human Visual Cortex and the Spatial Precision of Functional MRI. *Cerebral Cortex* 7:181-192.

FAUCONNIER, G. & TURNER, M.

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: University of Chicago Press.

GOODWIN, C.

1994 ‘Professional Vision’. *American Anthropologist* 96(3): 606-633.

1995 ‘Seing 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: University of Chicago Press.

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

HUTCHINS, E. & PALEN, L.
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 *Woman, Fire, and Dangerous Things*. Chicago: Chicago UP.

LATOUR, B.
1986 ‘Visualization and Cognition: Thinking with Eyes and Hands’. *Knowledge and Society* 6: 1-40

1987 *Science in Action*. Cambridge: Harvard UP.

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

McNEILL, D.
1995 *Hand and Mind: What Gesture reveal about Thought*. Chicago: University of Chicago Press.

McNEIL, D. (ed.)
2000 *Language and Gesture*. Cambridge: Cambridge UP.

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

SACKS, H., SCHEGLOFF, E. & JEFFERSON, G.
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., HADJIKHANI, N.K., MENDOLA, J.D., MARRETT, S, DALE, A.D.
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: 145-73.

This chapter is reprint of

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

CHAPTER 5

SEEING DIGITAL BRAIN IMAGES AS HYBRID OBJECTS TO WORK WITH

ABSTRACT

A significant part of MRI practice in neuroscience is spent in front of computer screens. The present essay describes ways in which neuroscientists, while looking at brain images, use their hands to gesture and manipulate digital screen in order to organize and make sense of their experimental data. Their embodied and ephemeral performances partake in the hybrid constitution of phenomenal objects articulated at the junction between the digital world of technology and the physical world of embodied social action. These sense-making processes concern the web of practical matters in the course of everyday laboratory work while they participate in the shaping of scientific facts.

1. INTRODUCTION

One of the crucial components of today's *visual and visualized experience* (Mirzoeff, 1999) is scientific visualization:

> Scientific visualization has the power to make visible and comprehensible atomic, cosmic, and common three-dimensional phenomena (such as heat conduction in engines, airflow over winds, or ozone holes). Abstract-information visualization has the power to reveal patterns, clusters, gaps, or outliners in statistical data, stock-market trades, computer directories, or document collections. Overall, the bandwidth of information presentation is potentially higher in the visual domain than it is for media reaching any of the other senses (Schneiderman, 1998: 522).

A key modern digital imaging system used for scientific and medical purposes is Magnetic Resonance Imaging (MRI) technology. The goal of MRI is to provide detailed images of the anatomical structure of internal body parts, such as 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 brain tissues are induced to emit a signal that is detected by the computer, where the signals, represented as numerical data, are converted into visual representations of the brain.

These *visual representations of what is not visual* are the concern of the present essay. But rather than assessing their numeric vs. visual character (see Alac, 2004), the focus is directed at the specific ways in which such "pictures of numbers" (Tufte, 1983) feature in work and collaboration (Heath & Button, 2002; Luff, Hindmarsh, Heath, 2000; Suchman, 1987). The essay discusses how pictures of numbers, despite their visual power, become *visible* (Goodwin, 1994) for MRI practitioners not only through vision, but also through the involvement of "the other senses" (Schneiderman, ibid.), so that they can be articulated as constituents of everyday laboratory dealings. By looking at instances of interaction among MRI researchers involved in the practical problem solving, I argue that the *meaning* of MRI images arises against the everyday work practices that the *interpretations* (Peirce, 1934-48) of the images presuppose and articulate. Those interpretations are ongoing, local processes that arise through coordination of bodies and technology in the environment of scientific practice.

Several recent projects in *science and technology studies* have been dedicated to the examination of social and cultural aspects of brain mapping. Anne Beaulieu (2002; 2004), Joseph Dumit (2004), and Andreas Roepeshort (2001; 2003), to mention a few, have pointed out that a full understanding of the visual phenomena in MRI and PET (Positron Emission Tomography) would, in addition to knowledge of neuroscience, require an account of the environment of production and reception of brain images. But while the work of these researchers, largely grounded in analysis of discourse on MRI, successfully engages with questions concerning objectivity, standardization and generalization, it remains somewhat removed from particular circumstances of scientific practice, and thus leaves questions of materiality and embodiment fairly unexplored. We still do not know about the material status that *digital brains* may conserve or acquire during particular instances of laboratory work in order for experimental data to be talked about, understood, and worked with. To address these types of questions, while staying attentive to discursive constructions, the present essay examines how embodied humans interact with the material conditions in which they act (Hayles, 1999: 195).

The examination is based on an ethnographic study conducted in three laboratories that research cognitive processes through employment of functional MRI[36]. To bring the interface between the body and technology (Ihde: 2002) to the forefront of inquiry, the traditional long-term participant observation is combined with micro-analysis of specific occurrences of events and practices (e.g., Goodwin, 2000a). The micro-analysis is performed on digital video recordings[37] to identify and track the multimodal semiotic forms, such as talk, gesture, gaze and body orientation, which scientists bring into coordination with digital screens in the process of knowledge construction.

The instances that the present essay deals with are two excerpts from the interactional and practical activity. The excerpts are representative of the corpus of thirteen

---

[36] In addition to the production of static images of interior regions of the body, 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 during that task.

[37] The study employed additional data collection methods, such as direct observation, semi-structured interviews, and analysis of documents like scientific papers, laboratory manuals, and scientific correspondence.

hours of video recordings collected during approximately ten months of fieldwork[38]. The interaction takes place between two neuroscientists involved in the identification of an MRI image artifact. The analysis of the excerpts, grounded in *practice-based theory of knowledge and action* (see Goodwin, 1994: 606), describes how scientists organize their looking through an involvement of multimodal semiotic modalities. The multimodal interaction suggests that specific features of digital images are first brought to sight in the social space of action. Rather than assuming that our understanding of digital images and the reality they represent is achieved exclusively through a sense of vision, the analysis suggests that the shared acts of multimodal interpretation, grounded in the context of ongoing everyday activity, function as processes that generate seeing/understanding.

To reconcile the observations of interaction with the questions of understanding and interpretation, the present study adopts some of the insights from *conceptual integration theory* (Fauconnier and Turner, 1998, 2002). Similar to the *distributed cognition* of Edwin Hutchins (1995), which describes the activity of socio-technical systems by using terms usually applied to internal cognitive processes, this study applies insights from *conceptual integration theory* to the workings of multimodal semiotic interaction. In such a way, it avoids any recourse to exclusively private mental phenomena (such as goals, intentions, purposes or consciousness) while it underlines the direct link between sense-making, inference, and communication (Peirce, 1934-48).

By focusing on the ordinary methods that scientists apply in their work practice, the essay aims to resume the discourse initiated by *laboratory studies* more than two decades ago (Lynch, 1985; Knorr-Cetina, 1981; Latour & Woolgar, 1979). As Latour & Woolgar (ibid.) stated regarding their ethnography of the neuroscience laboratory, the present essay is “concerned with the social construction of scientific knowledge in so far as this draws attention to the process by which scientists make sense of their observations” (pp. 32); the aim is an “examination of the methodical way in which observations are experienced and organized so

---

[38] Because of my membership in the cognitive science community, it is difficult to quantify that which

that sense can be made of them" (37). But, in addition to the study of language and graphical representations (see also Lynch & Woolgar, 1990), the present essay is primarily concerned with dynamic, embodied, and often non-representational performances and their interaction with digital screens. The study of such forms is broadly grounded in the *interpretative semiotics* of C.S. Peirce (ibid.), with its emphasis on the in-built dynamism of signs - where each sign always generates further signs and functions as a manifestation of a general phenomenon. Rather than "reconfiguring local expressions into various decontextualized representations" (Lynch, 1993: 310), the recourse to general postulates of interpretative semiotics (as well as borrowings from *conceptual integration theory*) allows for an engagement with the "messy" and "disordered" world of laboratory work to be *recovered*. Thus, semiotic performances are not *delocalized* to the *analytically transparent* networks of signs (ibid.: 292), but feature in the phenomenal field of laboratory praxis. Their identification is itself a situated activity that relies on specific details of practical action (Suchman, 2000:17).

2. SEEING SHEARING

The laboratories in which my fieldwork took place -'Laboratory I' and 'Laboratory II'- study human visual perception, mainly through MRI techniques. The primary interest of the two laboratories is to localize visual processes in specific brain regions, and to determine ways in which visual stimuli are processed there. Dissatisfaction with commercial *off-the-shelf* data analysis tools usually used in MRI data processing led members of both laboratories to design their own software. So designed software produces *inflated* and *flattened cortical maps*, and allows for more precision in identifying the location of brain activation. Because of the similarities in the methods practiced in the two laboratories, practitioners from both were interested in comparing the two software programs. The comparison was undertaken through the shared work and interaction between the members of the laboratories.

---

constitutes my fieldwork.

The present essay focuses on two short videotaped excerpts from those encounters. The excerpts come from an early stage of MRI data analysis where functional images (low resolution images that represent cognitive processes) and structural images (high resolution images that reveal the anatomy of the brain) have to be aligned with each other (see Figure 5-1). The interaction took place between the principal investigator (PI) from Laboratory I and a post graduate student (PD) visiting from Laboratory II. In order to carry out the software comparison, the two practitioners used software designed in Laboratory I to process experimental data previously analyzed with software used in Laboratory II.

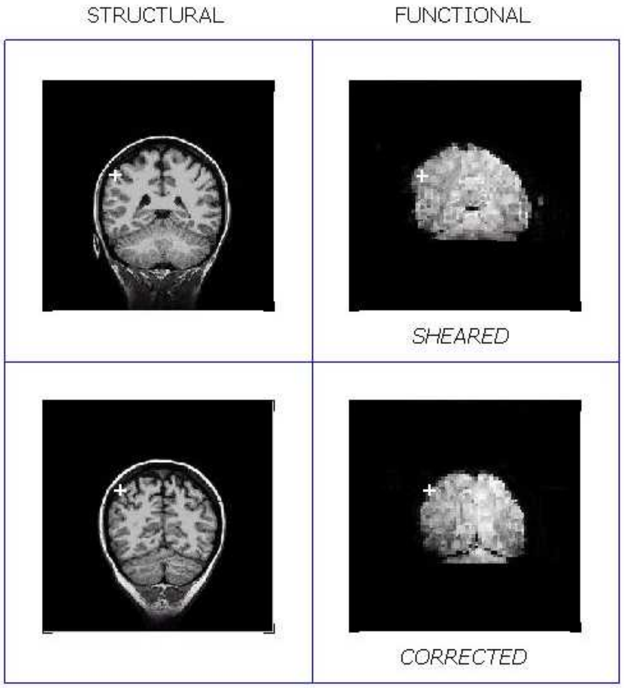


FIGURE 5-1: Functional and structural images as they appear on the computer screen

During the moments of collaborative work reported here, the PI and PD are seated in front of a computer in Laboratory I. A small functional image is displayed on the center of the screen. The use of mouse commands allows the PI to alternate the display of functional and structural images. Rapid switching between the images makes the difference between the two potentially apparent. Once spotted, the difference is discussed in terms of the inadequacies of a single image. The PI talks about a “shearing” artifact in the functional image. Of note is that the artifact was not noticed when the data were analyzed in the other laboratory.

As previously pointed out (e.g., Latour & Woolgar, ibid., Lynch, ibid.), work in elimination and purification of the data from artifacts is one of the central aspects of laboratory practice. In brain mappers’ jargon, “shearing correction” indicates a group of computational processes performed on images in order to correct a particular type of distortion. Similarly, the distorted image is described as being “sheared”. While trying to correct the image distortion, the PI demonstrates to the PD how an image can be classified as “sheared”, and shows how its correction is achieved through the software used in his laboratory. As the excerpts from the interaction testify, the two aspects of the activity are intertwined: the correction itself functions as a part of the classification process, and the classification - oriented towards the practical dealings - anticipates such processes through its own unfolding. An important issue then is to decipher what those practical doings entail when the *stuff* that the scientists are dealing with is *digital*.

The work with the digital stuff concerns the web of practices that go beyond the digital screen. The moments of noticing shearing in the PI’s laboratory vs. overlooking of artifact in the PD’s laboratory have to do with historical production of local actions and disciplinary-specific practical dealings. Those dealings, heavily invested in the enhanced ways of looking at the human body (e.g., Cartwright, 1995), shape and are shaped by the design, development, and use of instruments and technology. For example, the fact that Laboratory II’s “predecessor” laboratory was conducting research in human psychology rather than physiology and that the software developed there was originally designed as specific for function localization rather than brain mapping are significant for the course of the subsequent work. These factors orient the designers to build MRI software that directs the user to compute parts of the process automatically, not allowing her/him to engage with some of the components of the data manipulation process. Frequently, so designed software permits the user to look only at what neuroscientists call “regions of interest” rather than allowing her/him to inspect images of the whole brain. It is to be expected that this type of software and the associated techniques for looking at MRI data contribute to the local overlooking of image artifacts.

Importantly, the local dealings with digital images are not simply *shaped* by and do not merely *reflect* the dynamics of the encompassing socio-technical network. They are directly involved in the paths of scientific inquiry and the construction of scientific evidence. The PI proudly reported that the sessions of collaborative data analysis between the members of the two laboratories led Laboratory II to modify its software so that the inadequacies in the brain images can be “more easily spotted”. In other words, the involvement of semiotic processes in the socially shared space of image viewing influenced the configurations of technology and migration of techniques from one laboratory to the other, generating in turn effects on the future appearance of MRI representations.

The local and everyday matter of science, such as the moments of the PI and PD shearing correction, reveals how the ephemeral, transient and adaptable semiotic techniques feature in the interweaving between human actions and technology. By looking at those sites of interaction, the essay is concerned with the conceptual organization of setting in the context of technology use. In particular, it asks how the multimodal encounters with the digital world and the ways in which practitioners dynamically organize their working environments for understanding put in question our assumptions on how we carve and act upon our world.

2.1 PERFORMING SHEARED IMAGES

The first excerpt shows how the PI, as he notices that the data were not corrected for shearing in Laboratory II (he does so by using mouse commands to manipulate the computer screen so that serially organized images can be compared), sets to make the distortion of the data publicly visible. While doing so he ratifies the rational character of the procedures adopted and developed in his laboratory:

EXCERPT 1[39]

1 PI So (you) usually (0.1) at this point (0.5)=

2 I usually do a little bit of shearing (0.1)

3 to help get the: ((In lines 1-3 the PI alternates the display of MRI images on the computer screen))

4 you can see=

5 it's sort of [gesture exemplifying distortion] ((the PI looks toward the PD)

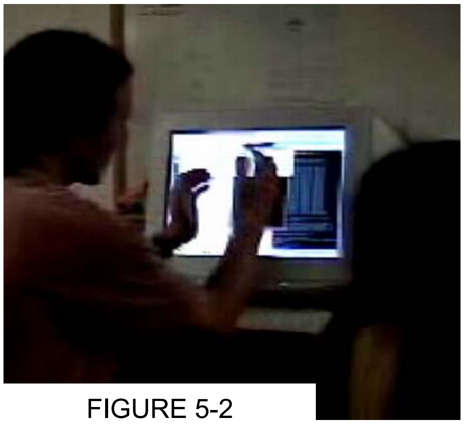


FIGURE 5-2

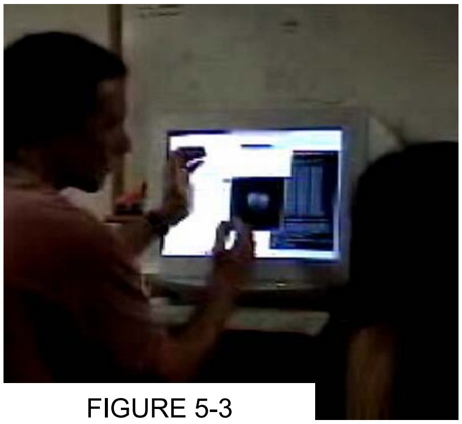


FIGURE 5-3

6 PD ((the PD affirms, but what she said exactly cannot be identified))

7 PI It's /sort of= ((the PI looks toward the screen))

8 PD /Yeah

9 PI it's actually sheared=((while he manipulates the computer screen with his right hand, PI's left hand is still in the position assumed in the previous gesture))

---

[39] Transcription conventions follow Goodwin (1994).

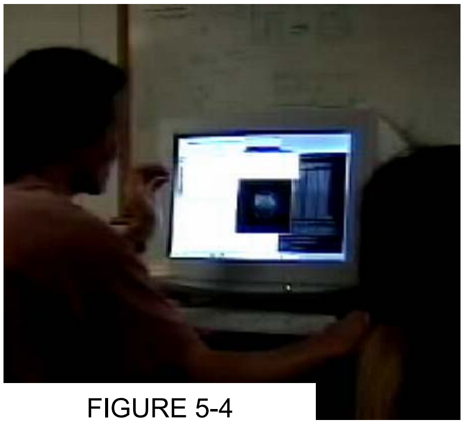


FIGURE 5-4

10 it's going up this way [the PI places his left hand, holding his thumb and index finger an inch or so apart, onto the screen, moves it along the border of the brain slice representation, and by clicking with his right hand on the mouse, changes the visual display from functional to structural image]

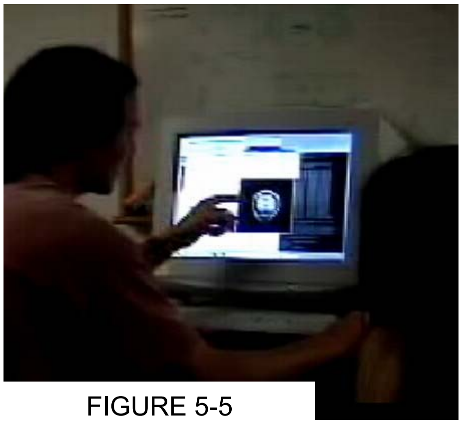


FIGURE 5-5

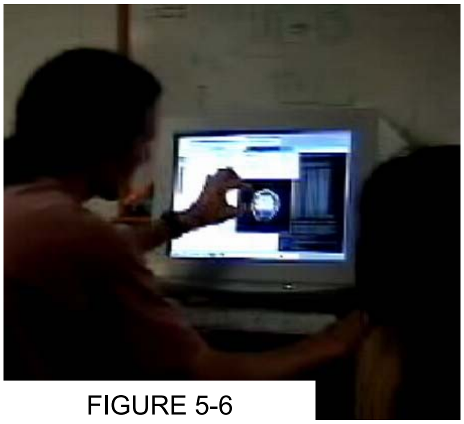


FIGURE 5-6

11PD Right ((the PD moves her upper body back, and then returns to the initial position))

The PI premises the construction of seeing by saying: "So (you) usually (0.1) at this point (0.5)=I usually do a little bit of shearing (0.1)". His usage of "usually" and "at this point" evokes the sequential order of the laboratory procedures and locates the specific instance of seeing, premised by the utterance, in such an order. Notice the interchangeable use of *I* and *you* in "so (you) usually (0.1) at this point (0.5)" (line 1), and "I usually do a little bit of shearing (0.1)" (line 2) reflecting the close coupling between *what is usually done* (I and the members of my laboratory) and *what should be done* (you). At the specific moment in time ("at this point"), defined in respect to the completion of the preceding steps of data analysis, the PI orients the joint attention toward the "shearing" procedures. He uses the mouse commands to alternate images on the computer screen while so searching for their defects (line 3).

An important component of the PI's activity is performed by the multimodal semiotic action. The utterance "you can see" in line 3, rather than just representing the PI's belief that the PD can already see the distortion in the image, prepares the terrain for the collaborative constitution of the visual scene. The PD is invited to participate in the performance where, through the PI's *professional vision* (Goodwin, 1994) and the techniques used in his laboratory, she will progressively be subsumed under the generic "you" that can see.

By marking the change in the discursive mode from the description of general procedures (lines 1-3) to orientation at particular local event, "you can see" (line 3) organizes the direction of the collective attention towards the computer screen. A similar function is accomplished by the PI's body movement in line 4. The PD can read the PI's orientation towards the screen as an invitation to orient her gaze toward the same site. The PD recognizes the function of the PI's semiotic action: she looks toward the digital image without yet acknowledging that she can see the distortion. Additionally, the two semiotic modalities – the talk and the body movement - while organizing the participation in the joint tuning of attention, delineate a space in which the image distortion will be made collectively visible. The PD will not learn to recognize the artifact just by looking at or manipulating the digital screen.

To make visible what he sees, the PI performs a "shearing" gesture (line 5). The PI places both of his hands in the space located between the two practitioners and the computer screen. He uses his hands to enclose a portion of the void space as if holding a round object (Figure 5-2). He then moves his left hand up while his right hand moves down (Figure 5-3). Throughout the process the two hands conserve the shape in which they were placed at the beginning of the gestural performance.

After performing the shearing gesture, the PI turns towards his addressee (line 5). In line 6 the PD signals her understanding and the co-participation in the action. During his turning toward the PD, the PI's hand remains in the position assumed during the previous gesture (line 9) (Figure 5-4). The steady hand position is significant - it allows for the linking of the enactment carried out in line 5 with the subsequent performance (line 10).

The linkage between the two units of action, the already executed and the future semiotic performance, is further developed through the PI's pointing gesture. When the PI directs his left hand toward the computer screen, he briefly points with his index finger to the upper left portion of the image (line 10) (Figure 5-5). After pointing to the image, the PI performs another shearing gesture. He places his thumb onto the screen, and, while holding his thumb and index finger an inch or so apart, he carefully moves his hand across the screen (Figure 5-6). During the execution of the gesture, he keeps his right hand on the computer mouse to alter the images on the screen. The sequence of knowledge construction is an instance of laboratory work. In both activities practitioners' hands and eyes play a central role.

2.1.1 CONSTRUCTING SEEING THROUGH MULTIMODAL ARTICULATION OF THE SOCIAL SPACE

Particularly salient elements of the interaction are the two gestures executed in lines 5 and 10. The gestures take part in the production of visibility by selecting and giving prominence to the relevant features of the images (Goodwin, 1994) – they categorize the images in terms of their shearing. They do so through their involvement with other multimodal semiotic resources where, in the context of laboratory practice, multiple *semiotic fields* (Goodwin, 2000b), such as the computer screen and the practitioners' bodies, are organized in spatial arrays. The superimposing and intertwining of the fields participates in the activity of making the experimental data visible. In this sense, the gesture and other dynamic, embodied semiotic means function as *instruments for organizing perception* (Lynch, 1990).

The first gesture - the gesture performed in line 5 - makes the shearing visible while being located at a relative distance, but in front of the brain representation. Whereas the digital image provides a rich substrate for the action and allows for the potentially numerous interpretational paths, the gesture functions as a highlighting mark that simplifies what would require a relatively complex search through the rich visual field. At the same time the practitioners' orientation toward the screen, their talk, the alignment with the digital images, and the general context of the practice constrain the interpretation of the gesture. Even though the gestural performance does not take place in synchrony with the PI's talk (in line 5 the PI performs the gesture after saying "it's sort of"), the PI's orientation toward the screen (lines 1-4), his saying "you can see" (line 4), his talk that follows the gesture – e.g., "it's actually sheared" (line 9), embedded in the context of the practical problem solving, indicate that the gesture is about the feature of the images.

The same is true for the gesture performed in line 10 (Figure 5-5): the gesture is interpretable in respect to the ongoing discursive activity and the laboratory setting in which it is lodged. Here, however, the visibility of the shearing is generated through a physical coupling between the gestural performance and the digital images. The gesture selects the salient feature in the images by touching and moving across the computer screen.

The way in which the images are aligned with the gestures, body orientation, gaze, and talk suggest that the PD *learns to see the shearing* not only by looking at the images, but also by participating in the multimodal semiotic structuring of the social space that surrounds the two practitioners. The process of learning and understanding is one of participation: it situates the practitioners within the visual field, not outside it.

2.1.2 DYNAMIC PHENOMENAL OBJECTS AS ARCHITECTURE FOR SEEING

In addition to their spatial organization, the observed multimodal semiotic assemblies have a salient temporal dimension. In line 5 the gesturing hands *perform* a distorted round object. But they are not directly brought into the form *representing* a distorted round object. Instead, they perform a process of transformation of a round object. Likewise, in line 10 they do not only index a feature in the image. Rather than just pointing to or representing the relevant feature of the image, the gestures participate in the act of interpretation given as an embodied enactment of the process of change through time. The dynamicity of the performance provides an account of how the distortion came about. The gestures generate explanations of the distortion in terms of a three dimensional (3D) object that is vertically *sheared*.

The gesture in line 5, coupled with the talk, body orientation and the digital image, is a dynamic, temporal performance of shearing that involves both of the PI's hands. The gesture in line 10, on other hand, is performed with just one hand. The PI uses his left hand to first touch the screen and then transforms it into a trajectory performed across the screen. The interaction between the manipulation of the computer screen and the gestural inscription of the movement reveals a direct involvement of the gesture in the task at hand. The gesture is not only used to show the features of the images to the PD. Rather, the gesture is implicated in the PI's coordination and comparison of the images to identify the discrepancies among them and thus *correct* the shared image. While he searches through the image series, the PI uses the gesturing hand to link the appearance of the cross (a sign that marks the corresponding spot in the series of brain representations, see Figure 5-1) across multiple images. This allows him to enact an *external memory trace* and single out the exact location of the artifact. The multimodal construction of visibility is a part of the local act of problem solving and data analysis.

While the PI moves his left hand across the screen and employs his right hand to alter the images and search for the distortion, the rapid alteration of the images generates appearance of motion (i.e., *apparent motion or PHI phenomenon*)[40] so that the discrepancies

---

[40] Since the experimental work of M. Werthaimer (1912) and A. Korte (1915) it is well known that when presented with a succession of still images, flashed at specific temporal and spatial distances, the observer will sometimes report seeing motion between the stimuli. Today, to study how such a

among the images can be understood in terms of a unified whole. The dynamic gestural performance and the apparent motion created as a part of the work done on the images are coupled with PI's utterance: "it's going up this way" (line 10). The utterance is an expression of what proponents of cognitive semantics call *fictive motion* (e.g., Talmy, 2000). Like the gesture, the utterance organizes the seeing in terms of motion and change even though the described entity – the fMRI image – is static. In such a way the multilayered performance - where the PI's talk *evokes* the motion, the manipulation of the screen *generates appearance* of motion, and the gesture *performs* such a phenomenon in the environment of practice - makes the motion a concrete, publicly available process which articulates a phenomenal object that changes through time.

Importantly, the dynamic embodied account of how the distortion came about implies that the process may be reversed by applying vertical shearing in the opposite direction. As the section 3 will illustrate in detail, the action of applying shearing in the opposite direction is in fact the process through which practitioners understand the work that needs to be done to correct the distortion. In this respect, the multimodal semiotic resources not only enact what happened in the past, but evoke the future practical actions.

Inasmuch as they are about the future practical dealings, the two sharing gestures and their interaction with the brain images are reminiscent of the Heidegger's distinction between things 'present-at-hand' (*Vorhandene*) and things 'ready-to-hand' (*Zuhandene*). The two practitioners do not represent sheared features as merely occurring objects that can be passively contemplated. Rather, the features are enacted as manipulable things that "subordinate themselves to the manifold assignments of the 'in-order-to'" (Heidegger, 1962: (69) 98). As lines 5 and 10 testify, the relevant feature of the MRI scans is first made manifest

---

phenomenon produces perception of motion neuroscientists employ fMRI. Muckli et al. (2002), for example, showed that the perception of apparent motion can be correlated with activations in hMT+(V5), the human motion complex that includes the human homolog of monkey's visual area MT and its satellites. The trick of apparent motion is explored by a variety of devices and techniques, from optical toys (such as the 18th century zoetrope), to contemporary cinema, television, and computer animation.

through the activity performed in terms of a direct engagement. The sheared feature becomes visible as a thing encountered through *circumspection* - a purposeful way of acting without the necessity of having a purpose in mind (see Dreyfus, 1991) - instead of a mere “natural object” positioned in front of the viewer to be seen. However, rather than allowing the images to fade away through such interaction (as the concept of *ready-to-hand* is often referred to in the literature on interaction with technology, see for example Dourish, 2001), the performance of the sheared object generates visibility of the images. Such visibility, in turn, provides practitioners with a sense of a concrete engagement.

And what the practitioners need to engage with are not the MRI images as the material brain slices that they depict, but the images as the digital MRI data that they index. Obviously, the shearing artifact doesn’t show that the scanned physical brain was “sheared”. Rather, the artifact is about a computational process that happened at the level of experimental data collected during the scanning procedure. Thus, the production of seeing through the performances of manipulable phenomenal objects points out that the encounter with the digital scans is not only about the visual into which the digital was *translated* – the digital data are not simply transformed into something else, estranged from its point of departure. Instead, the visibility is constructed so that it regards practical dealings, and what the practitioners need to deal with are the invisible numeric data.

As the opening quote from Schneiderman (ibid.) indicates, research in human-computer interaction (HCI) has pointed out that because of their aptness for understanding, the products of scientific visualization - such as MRI images - provide the human user with a significant power in handling the abstract data[41]. Rather than dealing with an enormous amount of numbers, digital images allow scientists to solve the abstract task in the visual realm. But, importantly, MRI images are visual versions of what is not visual inasmuch as they allow the practitioners to *deal* with the invisible. The excerpts from the interaction indicate how

---

[41] Neuroscientists themselves often point out that the largest part of the human visual cortex is dedicated to processing visual information.

the laboratory setting and the human encounters therein provide the practitioners not only with an abstract sense of vision, but with a seeing that involves the eyes as well as the hands. In addition to the brain images, the PI and PD use embodied multimodal semiotic means to share the meaning of the abstract data in terms of everyday objects and concrete actions.

Such a process, while making the digital data "accessible" and the brain images visible, generates something that does not exist neither at the level of digital data nor brain images taken in isolation. As already pointed out, the multimodal interaction in lines 5 and 10 gives the abstract cause in terms of motion or change. Motion is one of the central features of our world and our action in the world. While practitioners are aware that motion is not at the origin of the shearing artifact, their invoking of motion allows them to deal with the task at hand in terms of the everyday coping in the world of objects that change through time. In other words, the practitioners know that the "shearing" of the image is a consequence of an abstract computational transformation. Even so, their semiotic performances indicate concrete, physical transformations. In other words, practitioners' practical accounts of the experimental data are given in terms of objects that can be seen, experienced and dealt with, rather than in terms of causal and abstract explanations of the artifact.

3. MAKING THE DEALINGS VISIBLE

After making the artifact publicly available through the performances of dynamic objects - the *in-order-to* structures that show themselves only in dealings - the work in purification of the experimental data needs to be explicated in a step-by-step fashion. To access the data by making the work visible, the practitioners deal with MRI images – the digital structures that generate effects of manipulability.

EXCERPT 2

12 PI You rotate (into the) [PI places right hand on the screen as if holding the object represented, and rotates it towards the left]

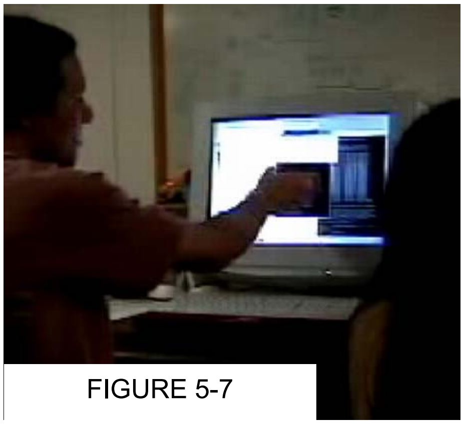


FIGURE 5-7

13 so that you know [shakes the hand while keeping it in the same position]

14 (so that) the expanded part is sort of= [indicates up with his index finger]

15 in one of the axes [gestures as if drawing two axes]

16 and you squish [gestures as if squeezing the object with his right hand]

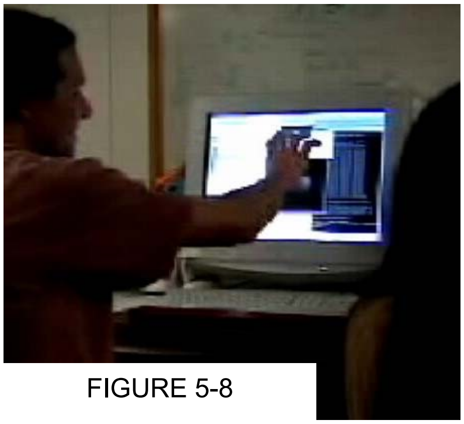


FIGURE 5-8

17 and then (you) rotate back [rotates the hand holding the imaginary object towards the left]

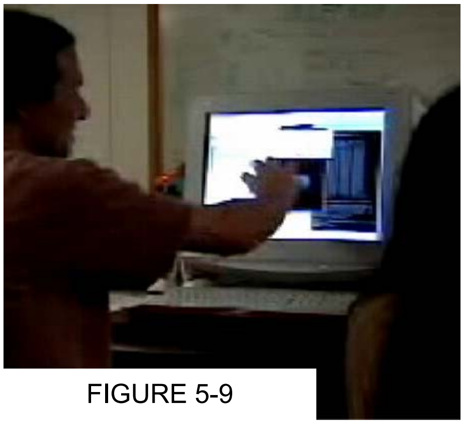


FIGURE 5-9

18 and un-squish [quickly moves hand as if mimicking the expansion of the representation]

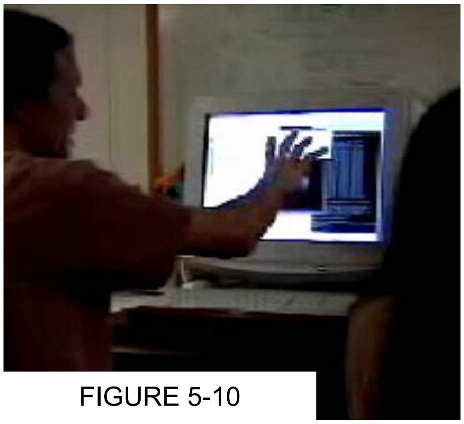


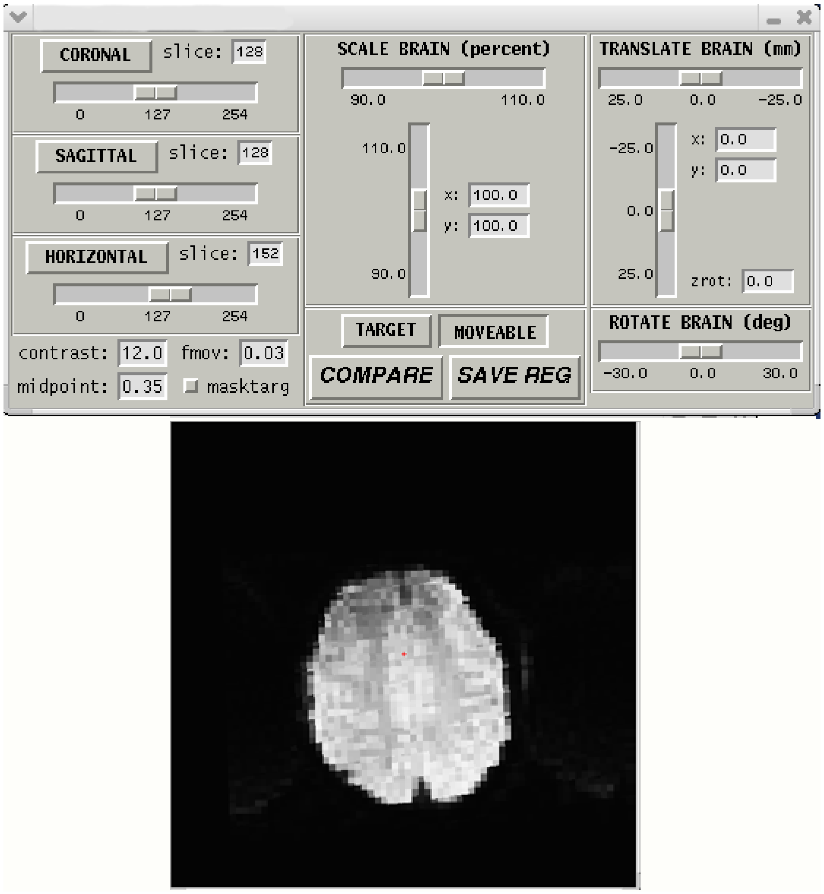
FIGURE 5-10

19 so: rotate 30 degrees, shrink vertically, rotate back, stretch vertically [turns towards the PD and smiles]

20 PD: Uh-huh

The organization of the activity reported here is shaped by the way in which the software program designed in Laboratory I works. Because of the program's limitation, the MRI image series cannot be automatically corrected. Instead, to achieve the correction announced by the gestures in lines 5 and 10 (Excerpt 1), the series has to go through a sequence of transformations.

Figure 5-11 shows a detail of the computer screen during the activity of shearing correction. The user needs to interact with the scroll bars via mouse commands to direct mathematical transformations on the digital data. The transformations in the digital data are indicated in the visual format: in response to the manipulations of the command menu, the brain images appear modified.



CORONAL slice: 128

SAGITTAL slice: 128

HORIZONTAL slice: 152

contrast: 12.0 fmov: 0.03

midpoint: 0.35

SCALE BRAIN x: 100.0 y: 100.0

TRANSLATE BRAIN x: 0.0 y: 0.0 zrot: 0.0

FIGURE 5-11: Detail of the computer screen during the shearing correction. (Proportions have been modified to enhance the visibility of the menu)

Once the shearing artifact is detected in the images, four consecutive commands need to be selected. First, the scroll bar labeled “ROTATE BRAIN (deg)” needs to be “moved” towards the right. When this is done, the vertical scroll bar labeled “SCALE BRAIN (percent)” has to be moved down. The third command is the movement of the scroll bar “ROTATE BRAIN (deg)” towards the left. And finally the horizontal scroll bar “SCALE BRAIN (percent)” has to be moved towards the right. The commands’ function is illustrated in Figure 5-12. To make the distortion more easily perceived, the Figure represents the brain image in a schematized rectangular rather than a round form.

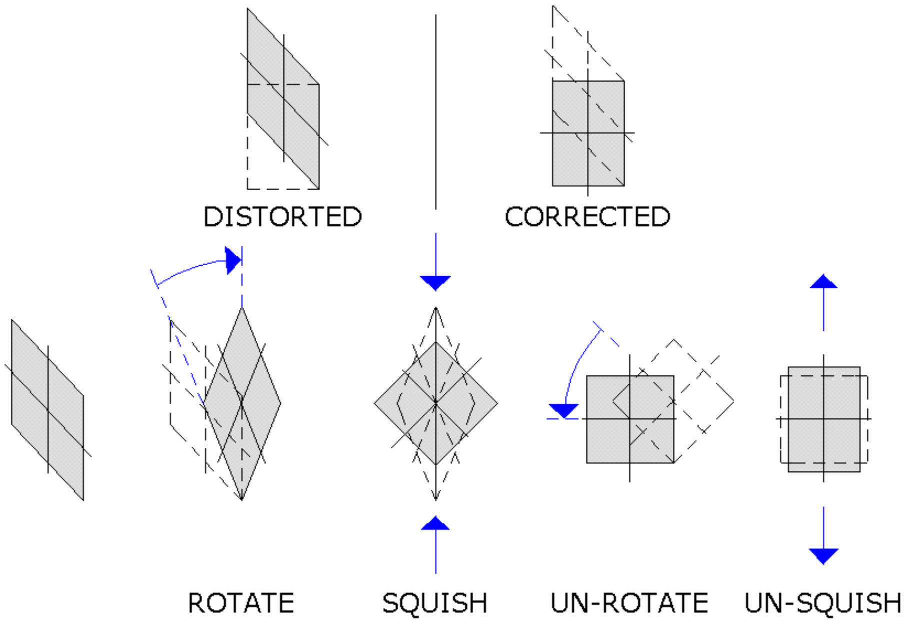


FIGURE 5-12: Schematized representation of the four stages of moving correction

But, like the brain images on the screen, the figure is a partial, two-dimensional (2D) representation of the experimental data.

Despite the appearance of the 2D representations on the computer screen, the practitioners need to work with 3D data. They need to perform the changes on something that is more akin to a series of 2D images than to a single 2D representation. The procedure requires significant expertise. The practitioners need to infer the shape of the 3D object before and after its digital modification while looking at a 2D MRI image. To do so, the design of the computer interface allows them to work with the scroll bars and labels evoking the physical manipulation of 3D objects - e.g., “rotate brain”. Of note is that these features only evoke objects and actions; the environment in which the interaction between the user and the experimental data takes place is a 2D digital space.

But, as the video-taped record indicates, to deal with the task in a more immediate way, the practitioners have another resource at their disposal: the social space of semiotic interaction. In line 12 the PI places his right hand on the computer screen and rotates it towards the right (Figure 5-7). Next, in lines 14-15, he explains that the rotation has to position the brain image so that its expanded part is located in a vertical position ("on one of the axes"). The PI points up (line 14) and gestures to draw an ephemeral representation of the two Cartesian axes (line 15). This position is crucial for the next step of the transformation ("squishing") where a vertical force is applied onto it. Lines 16-18 are dedicated to the illustration of the acts of "squishing" (Figure 5-8), "rotating back" (Figure 5-9), and "un-squishing" (Figure 5-10). When the four-step transformation has been performed, the PI turns towards the PD to verify her participation and understanding of the action. The PI's bodily movement and the PD's affirmation ("Uh-huh", line 20) mark completion of the unit of meaning-making.

An analogous activity was recorded in the laboratory two months earlier, during another instance of data analysis. The PI and two graduate students were seated in front of the computer screen when the PI noticed a distortion in the MRI image. While orienting toward the computer screen, the PI described how the distortion should be corrected:

EXCERPT 3

21 PI If you rotate 30 degrees this way [places right hand on the screen as if holding the object represented; rotates it approx. 30 degrees to the left]

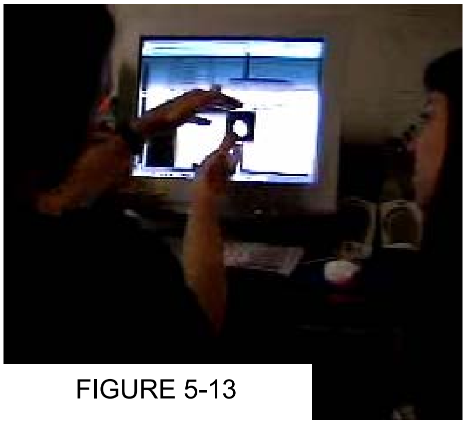


FIGURE 5-13

22 and then squash [places both hands on the screen, as if holding the object represented and slightly pushing it from both sides]

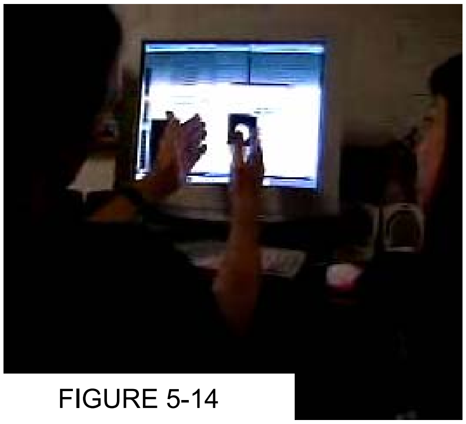


FIGURE 5-14

23 and then un-rotate [rotates both hands holding the imaginary object towards the left]

24 and un-squash [quickly moves right hand, as if mimicking the expansion of the representation]

The semiotic action performed here is quite similar to the one reported in the preceding excerpt where the PI gestures with one rather than both hands. The rituality of the performance reveals the tight connection between semiotic enactments and the world of instruments and technologies – the design of the digital screen, for example, and “usual” laboratory activities - which themselves are always bound to local contextures of relevancies (Lynch, 1993). At the same time it also suggests the pervasiveness of the gesture in the production of seeing. The PI reaches toward the computer screen and acts as if he were holding a round 3D object and moving it towards the left/right (lines 21 and 23) (Figure 5-13). Similarly, he performs the action of “squashing” over the image as if he were squeezing a 3D object (line 21) (Figure 5-14). By performing the physical manipulation on the MRI data, the gesture participates in the production of visibility by making the encounter with the phenomenal object possible. Whereas in lines 5 and 10 (Excerpt 1) the two shearing gestures participate in the *performance of ready-to-hand objects* so that the future work on such objects can be easily grasped, here the gestures participate in the performance of the exact steps of such a work. Thus, the interaction renders the future work not only graspable, but directly visible.

3.1 BLURRING THE BOUNDARY BETWEEN THE DIGITAL WORLD AND THE WORLD OF PRACTICE

The practitioners know that the shearing correction needs to be accomplished through the mathematical manipulation of digital data. Nonetheless, the abstract computational processes - not accessible through direct inspection, but actually performed on the digital data – are largely delegated to the machine. While they consider the numerical data and the mathematical processes to be the central and ever-present objects of their practice (see Beaulieu, 2002), the practitioners do not work with them through a direct access.

On the contrary, the visual representations of the data play a prominent role in the practitioners’ dealings with the problem at hand. The PI searches for the expected type of distortion and identifies it in the MRI images. After selecting a transformation command, the visually represented brain slice changes its form, providing so the feedback on the command choice. In fact, because they are digital structures, rather than traditional “immutable” signs (Latour, 1986), the MRI images have a capacity to provide the practitioners with a link to the experimental data and their mathematical handlings.

But despite the PI and PD’s orientation toward the computer screen, not everything is, or even should be, available there. For example, the PI’s command selection produces consequences not only for the brain slice displayed on the screen, but also for the data representing the whole brain. Rather than exhibiting the sequentially organized images one by one, the PI employs the embodied semiotic modality to render the information on the entire data set publicly available. While the talk evokes not only what is invisible on the screen (i.e., the three-dimensionality of the brain representation), but also what has to be done with it (i.e., rotation, scaling, etc,), the gestures render such actions visually present[42].

Thus, to understand the process of making the data visible, the semiotic field of embodied multimodal action, in additional to the field of digital images (which for MRI researchers always implies the realm of abstract digital data) needs to be taken in account[43]. The two semiotic fields, when taken in isolation, are rather poor. The images are completely static - simply displayed in the position assumed *before* and *after* their computational manipulation. Similarly, the linguistic labels on the screen and the PI’s utterances describe motion, but nothing is moving on the screen. The sense of movement is generated through the manipulation of the screen and the PI’s gesturing hand that touches the screen. The gesturing hand is shaped as if holding a 3D object, but without its interaction with the other semiotic resources, the enacted object remains generic (i.e., there is nothing in the gesture that defines

---

[42] It may be suggested that this *problem* could be *solved* through a visual display that would reveal all the richness of the data – “it is only a matter of time and technology”. Regarding this point, the PI remarked that if in fact all the layers of the 3D data were displayed (while discussing this point he evoked an image of a *translucent cabbage*), their richness would overwhelm the user, leaving her/him unable to deal with the scene. On the contrary, by including the gestural action in the process the practitioner is present with a form that evokes the 3D character of the data, while leaving the overwhelming detail concealed.

the object as the brain). Importantly, though, the two excerpts suggest that the practitioners, to deal with the problem of shearing, do not employ the semiotic fields separately, or evoke them one at a time.

In this respect, the specific location of the gesture is significant. Its unfolding takes place in a close proximity to the computer screen. The gesturing hand, accurately positioned around the contours of the brain-slice representation, touches the digital screen while it performs the “rotation”, “squishing”, and “squeezing” of a round object. The bringing together of the talk, gesture and the structure in the environment (what Goodwin, 1995; 2000b calls *environmentally coupled gestures*) helps the observer to see the images as experimental data to be manipulated (Figures 5-7 – 5-10, 5-13, 14).

While the enactments of rotation, squishing, and squashing are about the processes carried on the digital data, they, however, are not representations of those processes - nothing like rotation, squishing, and squashing is performed on the data. Rather, the semiotic enactments, conceived as performances of practical dealings and tightly coordinated with the digital images, index the invisible layers of abstract work while delegating its mathematical transformations to the machine. The practitioners are thus enabled to grasp the nature of the distortion and its correction by encountering something that is simultaneously about digital and physical, abstract and concrete, human and machine, 2D and 3D, objects and actions, present and future[44].

So conceived hybrid phenomenon exhibits emergent properties that do not exist in any of the semiotic fields taken in isolation, but are feature of the integration among multiple domains. Even though the digital images and the experimental data that they index cannot be physically manipulated, the hybrid structure provides the experience of such actions. The centrality of such hybrid phenomenal objects in the process of acquisition of expertise

---

[43] This certainly does not deny that an *expert* doesn’t need to resort to such public resources. But it does point out the public grounding of expertise.

[44] Ochs, E., Gonzales,P., Jacoby, S. (1996) describe similar phenomena as *referentially ambiguous entities*.

suggests the distributed and situated character of professional knowledge (Hutchins, 1995; Suchman, 1987) even at the level of the minute instances of multimodal interaction.

The importance of hybrid semantic structures in human understanding has been pointed out by *Conceptual Integration Theory* (Fauconnier & Turner, 2003; Coulson, 2000). Rather than accounting for stable knowledge structures represented in long-term memory, the theory identifies systematic projections of language, imagery, and inferential form to model the dynamic evolution of speakers' on-line representations (Coulson & Oakley, 2000). To explain such phenomena, Fauconnier and Turner (1998) conceive of a set of general cognitive processes – *conceptual integration or blending* - used to combine conceptual structure and to posit *mental spaces* (Fauconnier, 1994) on which conceptual integration operates. Mental spaces are short-term mental scenarios integrated on-line in conceptual integration networks to create new meanings through the juxtaposition of familiar material. The prototypical networks consist of two input spaces as well as the blended space that reveals the emergent, integrated meaning.

The semiotic performances of shearing artifact share some features with the process of conceptual integration. First, the coordination between the multiple meaning domains brings certain features to attention while leaving others in the background. For example, the abstract processes being performed on the digital data are not brought into the foreground during the acts of practical understanding. At the same time, the elements central for dealings with the task at hand are made salient while being crisscrossed and compressed. The hybrid semiotic construct is at the same time about the brain image that indexes the experimental data, and the embodied multimodal performances that make such data and their manipulations publicly available. The mutual reference and integration provides access to the problem solution in terms of ordinary actions. The practitioners deal with the experimental data and their correction by performing round objects that share some visual features with the digital image and are being rotated, squished, and squashed, even though neither the digital data nor the MRI images can be rotated, squished, and squashed. The process of hybridization is largely generated in the social space of action. Rather than regarding exclusively mental phenomena internal to single individuals, as it is largely the case in the applications of the conceptual integration theory, important elements of the production of visibility happen in the public space of action, where they involve integrations between multiple semiotic fields generated through the use of gesture, digital images, and body orientation as features of the practical problem solving.

Edwin Hutchins (2005) pointed out the existence of conceptual integration networks where one input space is a mental space in the usual sense, while the other one has the structure of a physical object or event. Hutchins calls such phenomena *material anchors for conceptual blends*. Examples of such structures are watches, the method of loci, and people standing in a line. Hutchins' examples show how conceptual relations can be mapped onto relations between material elements. One of the emergent properties of such a construct is the stability of conceptual representation. Hutchins' proposal can be further developed in respect to multimodal semiotic actions.

The examples of shearing correction accomplished by the two practitioners suggest that multiple semantic spaces can be largely instantiated in the environment of practice as semiotic fields. To conceptualize the images and the physical manipulations of objects, practitioners can rely on digital displays and semiotic performances (while at the same time partially delegating and negotiating the process of the manipulation of digital data with the machine). Hutchins suggested that these external forms could ground the internal cognitive processes while providing them with precision. But while MRI images may function like Hutchins' material anchors, the semiotic performances provide the process with additional characteristics. What is happening in the external world of practice is enriched with a temporal dimension: the practitioners can observe the dynamic accounts of how shearing took place and how it can be corrected. Even more, the integration between the domains can be importantly developed in the shared world of practice. The PI touches the screen to coordinate the digital world with the domain of things and actions. While in Hutchins' type of networks the products of integration are still largely accomplished internally, the shearing examples show how the coordination between the domains can significantly happen in front of the actors' eyes, in the space of embodied, culturally organized practical actions.

3.2 FROM THE PERSPECTIVE OF PRACTICAL SKILLS

The description of the coordination between the multiple semiotic fields – the embodied semiotic performances, digital images, and data that they index - may suggest their autonomous existence. This impression is further sustained by the linguistic expressions such as "hybrid", "integrated", etc..

From the perspective of practical skills, however, it is not evident that the practitioners first deal with separate meaning domains, nor that their semiotic actions are plain translations of the digital world into the world of practice. The existence of the separate meaning domains may rather be derivative of the integrated domain. The PD understands the features of MRI data by first encountering "images that can be squished, rotated, etc.". The knowledge of the separate domains progressively develops through the encounter of the hybrid form within the world of practice: the PD learns to see the shearing artifact in the isolated MRI by participating in the activity of rotation, squishing and squashing.

Similarly, the interaction between the multiple meaning domains is not just an unidirectional process of mapping from the *source* (here the manipulation of objects) to the *target* domain (here the digital data) (Lakoff, 1987), where the source domain provides a metaphoric framing of the target domain. Because the shearing features are seen as something that requires manipulation, the semiotic actions are not only about the operations executed in the physical world of the 3D objects. Instead, akin to *conceptual integration theory*, they always also concern digital data[45]. In fact, in the world in which ordinary actors interact and choose practical courses of action, the hybrid semiotic structure allows them to encounter selected aspects of the problem, identified in respect to their immediate relevance and their understandability, all at once. The interaction with the compressed form generates a sense of direct engagement with the experimental data while providing a common-sense account of the laboratory procedures.

4. DISCUSSION

Jonathan Crary (1990), among others, has pointed out that the introduction of computer-generated imagery leads to something he calls "abstraction of the visual". He proposes that digital images, unlike traditional analog media such as film, photography and television, are not relative any longer to a point of view "located in the real space":

> Computer-aided design, synthetic holography, flight simulators, computer animation, robotic image recognition, ray tracing, texture mapping, motion control, virtual environment helmets, magnetic resonance imaging, and multispectral sensors are only a few of the techniques that are relocating vision to a plane severed from a human observer. … Most of the historically important functions of the human eye are being supplanted by practices in which visual images no longer have any reference to the position of an observer in a "real", optically perceived world (Crary, ibid.: 1-2).

While one might think that this "abstraction of the visual" and the difference between new digital and traditional analog media necessarily entail a disembodiment of the digital image user, the interaction in MRI laboratory reveals a somewhat different state of affairs. Digital brain images, coupled with body orientation, gesture, and talk, enable practitioners to read the experimental data in terms of embodied actions that take place at a level comparable to physical, real-world engagement. Rather than producing an estrangement of the direct engagement with objects, digital images enable practitioners to re-enter the world of culturally meaningful and embodied actions.

---

[45] For a comprehensive description or the differences between *conceptual metaphor theory* and *conceptual integration theory* see Grady, Oakley, & Coulson, 1999.

This essay argues that to gain meaning, MRI images are incorporated in the laboratory practice as part of a pattern of actions: the practitioners notice and share the structure of brain scans by *acting upon* and *manipulating* the relevant information. The interaction reconstitutes the visual scene in terms of "shearing" and its correction through rotation, squishing and squashing. In this sense, vision is an embodied process not only because the eye is an organ and visual processes are lodged in the brain, but also because the acts of seeing are produced through an interaction between purely visual information and the world of meaningful embodied actions. In other words, the visibility is not only relative to what goes inside the practitioners and to what is present on the screen. Seeing is tied to actions that arise out of our experiences in the manipulation of objects. What is more, seeing takes place at interindividual territory (Vološhinov, 1986), and is about the establishment of communal meanings.

There is a substantial body of literature that deals with the construction of vision as a historically structured, situated and interactionally organized phenomenon (e.g., Goodwin, 1994, 2000c, 2003; Heath & Hindmarsh, 2000; Hindmarsh & Heath, 2000; Suchman, 2000; Ueno, 2000; Nishizaka, 2000). Following this line of research, the present work discusses "seeing" as a process situated at the intersection between instruments and technology, practices, setting, and practitioners' embodied accounts. It points out how the local construction of seeing indicates past dealings and practical know-how, and how it receives its rhetorical force *in respect to* the "usual" procedures of laboratory members. These procedures are, furthermore, nested in larger, historically evolving, socio-technical networks whose specifications are bound to the locally instantiated assemblages of instruments, embodied techniques, and everyday discourses in the laboratory (e.g., Lynch, 1993). Why did the members of Laboratory II, but not Laboratory I overlook the artifactual feature of the brain image? And while exhibiting some elements of rituality (similar performances have been observed in Laboratory I on several occasions) - which certainly indicates the alignment of the semiotic performances with the "usual" character of laboratory procedures and historically structured digital displays – the construction of visibility is accomplished with regard to the local circumstances and contingencies, transient arrangements of bodies and instruments, which arise in dealing with the specific problem at hand (Heath & Hindmarsh, 2000; Suchman, 1987) – making the image artifact publicly visible.

In exposing the involvement of semiotic performances with practical problem solving, the essay implicitly argues for peircian interdependence between sense-making, inference, and communication. The ways in which the PI's hands participate in rendering the shearing visible while being directly involved in searching through and organizing the experimental data, expose the intersection between that which is internal and external, problem solving and communication, individual and collective (see also Alac & Hutchins, 2005).

But in particular the essay is concerned with the production of seeing through the enactment of dynamic, publicly shared phenomenal objects. The PI's gestures in the shearing example do not only function as indexical signs or transient inscriptions imposed on the computer screen to categorize and make it visible. They also participate in the process where the screen is taken in the construction of phenomenal objects so that the PD can learn to spot the image artifact in respect to what should be done, can be done, and is "usually" done in the PI's laboratory.

This *engaged seeing* does not entail a *translation* of the invisible layer of the numerical data into something that can be passively observed through an abstract sense of vision. Rather, to see the features of interest in the brain image, the PI and PD coordinate the gestures, bodies, and digital screens with the implied layer of abstract mathematical data. The coordination explores the boarder between the virtual and the physical, 2D and 3D, abstract and concrete: it generates emergent properties and concerns distribution of agency. The phenomenal objects provide the practitioners with the sense of rotation, squishing, or squashing, for example, even though nothing is being rotated, squished or squashed, and nobody is rotating, squishing or squashing anything. The PI and PD talk about, see and experience a round object that changes and is being changed through time while they interact with menus on a computer screen displaying 2D brain scans to direct mathematical transformations on the numerical data.

It has been often pointed out that involvement with the digital screen transposes the user into an alternative realm – a realm of fiction. To engage with the digital data in an immediate manner, the user needs to *project her/himself on the other side of the screen* or *pass through the screen* to *enter the virtual world of fiction* (e.g., Morse, 1998; Lister et al., 2003; see also Turkle, 1995)[46]. By analogy, it could be claimed that the MRI practitioners' embodied semiotic accounts are *imaginary, fictive* or *metaphoric* renderings instantiated in the public space of action.

Several recent studies have dealt with the problem of imagination as a public process accomplished through the coordination of participants' conduct and the material world of their practice (Murphy, 2004, 2005; Nishizaka, 2000, 2003; Suchman, 2000). In the examples reported by the authors, the participants treat what is imagined and performed in the interaction space - a loading dock, submarines' routes in a computer game, or a highway - as something that has a potential of a future existence: the loading dock could be constructed, the submarines could be visualized as taking an alternative route on the computer screen, the highway could be lowered and the earth could be removed. In the shearing example, however, that which is publicly performed is not conceived as imaginary nor is it seen as something that has a potential of an existence outside the local moments of its embodied enactment. In fact, we may not even be able to postulate a full capacity for *seeing in front of our mind's eye*[47] something that is 2D and 3D, abstract and concrete, digital and material.

At the same time, and despite its fantastic character, the practitioners treat the publicly performed phenomenal object as something real and ordinary. The ways their embodied accounts develop through the day-to-day work of science suggest that the fantastic hybrid

---

[46] For an example of the popular rendering of the idea see Walt Disney Productions science fiction film *Tron* directed by Steven Lisberger, 1982.

[47] Nishizaka, 2003 contains a comprehensive discussion and critique of the idea of mental image.

entities are nothing else but common-sense. While tacitly *asking* practitioners to partially delegate the work to the machine, the production and understanding of multifaceted, phenomenal objects provide practitioners with a sense of direct engagement. The negotiation and hybridization between the digital and the physical is a practical way of understanding and dealing with the world.

Dealing with the world in this manner - *performed* through multimodal semiotic interaction and its coupling with technology - renders the PI's seeing of the MRI data particularly powerful. Not only does the PI's reading of the images prevail through moments of local interaction; it also allows the embodied techniques and the methods of organizing the tools and instruments practiced by his research group to propagate and leave its traces in the ways in which scientific facts are structured beyond the walls of his laboratory.

ACKNOWLEDGEMENTS

I would like to thank the following people for their contribution to this project: Lucy Suchman, Monika Buscher, Lisa Cartwright, Aaron Cicourel, Seana Coulson, Hollie Crower, Alessandro Duranti, Ryan Downey, Gilles Fauconnier, Deborah Forster, Charles Goodwin, Ed Hutchins, Maurizio Marchetti, David Mindell, Maggie Mort, John Rooksby, Ayse Saygin, Marty Sereno, Dan Shapiro, Tim Koschmann, Patrizia Violi, Laura Watts, Hsin-Hao Yu, and the participants in the ethnographic study.

WORK CITED

ALAČ, M.
2004 “Negotiating Pictures of Numbers”, *Journal of Social Epistemology* 18 (2): 199-214.

ALAČ, M. & E. HUTCHINS
2004 "I See what you are Saying: Action as Cognition in fMRI Brain Mapping Practice", *Journal of Cognition and Culture* 4:3: 629-661.

BEAULIEU, A.
2002 “Images are not the (only) truth: Brain mapping, visual knowledge, and iconoclasm”, *Science, Technology, & Human Values* 27, 53-86.

2004 “From brain bank to database: the informational turn in the study of the brain”, *Studies in History and Philosophy of Biological and Biomedical Sciences* 35: 367-390.

CARTWRIGHT, L.
1995 *Screening the Body: Tracing Medicine’s Visual Culture*, University of Minnesota Press.

COULSON, S.
2000 *Semantic Leaps: Frame-shifting and Conceptual Blending in Meaning Construction*, Cambridge: Cambridge UP.

COULSON, S. & OAKLEY, T.
2000 “Blending Basics”, *Cognitive Linguistics* 11, 3-4: 175-196.

CRARY, J.
1990 *Techniques of the Observer*, Cambridge, Mass: MIT Press.

DOURISH, P.
2001 *Where the Action is: The Foundations of Embodied Interaction*, Cambridge, MA: MIT Press.

DREYFUS, H.L.
1991 *Being-in-the-world: A Commentary on Heidegger’s Being and Time, Division I*, Cambridge, MA: MIT Press.

DUMIT, J.
2004 *Picturing Personhood: Brain Scans and Biomedical Identity*, Princeton, NJ: Princeton UP.

FAUCONNIER, G.
1994 *Mental Spaces*, Cambridge: Cambridge UP.

FAUCONNIER, G. & TURNER, M.
1998 “Conceptual Integration Networks”, *Cognitive Science* 22: 133-187.

2002 *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*, Basic Books.

GARFINKEL, H.
1967 *Studies in Ethnomethodology*, Englewood Cliffs, NJ: Prentice-Hall.

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

1995 "Seeing in Depth", *Social Studies of Science* 25: 237-74.

1997 "The Blackness of Black: Color Categories as Situated Practice", In L. Resnick, R. Säljö, C. Pontecorvo and B. Burge (eds.), *Discourse, Tools and Reasoning: Essays on Situated Cognition*, New York: Springer-Verlag.

2000a "Practices of Seeing, Visual Analysis: An Ethnomethodological Approach", In T. van Leeuwen & C. Jewitt (eds.), *Handbook of Visual Analysis*, London: Sage.

2000b “Action and embodiment within situated human interaction”, *Journal of Pragmatics* 32: 1489- 1522.

2000c "Practices of Color Classification", *Mind, Culture and Activity* 7(1-2): 19-36.

2003 "The Semiotic Body in its Environment", In J. Coupland & R. Gwyn *Discourses of the Body*, New York: Palgrave/Macmillan.

GRADY, J., OAKLEY, T. & S.COULSON
1999 “Conceptual Blending and Metaphor”, In G. Steen & R. Gibbs (eds.) *Metaphor in Cognitive Linguistics*, Amsterdam and Philadelphia: John Benjamins.

HEIDEGGER, M.
1962(1927) *Being and Time*, New York: Harper.

HAYLES, K.N.
1999 *How we became Posthuman: Virtual Bodies in Cybernetics, Literature* and Informatics, Chicago: University of Chicago Press.

HEATH, C. & G. BUTTON (guest eds.)
2002 “Editorial Introduction”, *Special Issue on Workplace Studies, British Journal of Sociology* 53(2): 157-161.

HEATH, C. & J. HINDMARSH
2000 “Configuring Action in Objects: From mutual space to media space”, *Mind, Culture and Activity* 7(1-2): 81-104.

J. HINDMARSH & C. HEATH
2000 “Embodied Reference: A study of deixis in workplace interaction”, *Journal of Pragmatics* 32: 1855-1878.

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

2005 “Material anchors for conceptual blends”, *Journal of Pragmatics* 37/10: 1555-1577.

HUTCHIN, E. & PALEN, L.
1993 "Constructiong Meaning From Space, Gesture, and Talk", In L.B. Resnick et.al. (eds.) *Discourse Tools and Reasoning*, Springer Publishing.

IHDE, D.
2002 *Bodies in Technology*, London: University of Minnesota Press.

KNORR-CETINA, K.
1981 *The Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science*, Oxford: Pergamon Press.

KORTE A
1915 “Kinematoskopische Untersuchungen” (Kinematoscopic Investigations), *Zeitschrift für Psychologie* 72, 194-296.

LAKOFF, G.
1987 *Woman, Fire, and Dangerous Things*, Chicago: Chicago UP.

LATOUR, B.
1986 “Visualization and Cognition: Thinking with Eyes and Hands”, *Knowledge and Society* 6: 1-40.

1987 *Science in Action*, Cambridge, MA: Harvard UP.

1996 (1992) *Aramis: The Love of Technology*, Cambridge, MA: Harvard UP.

1987 *Science in Action*, Cambridge, MA: Harvard UP.

LATOUR, B. & WOOLGAR, S.
1979 *Laboratory Life. The Social Construction of Scientific Facts*, London and Beverly Hills: Sage.

LISTER, M., DOVEY, J., GIDDINGS, S., GRANT, I. & K. KELLY
2003 *New Media: A Critical Introduction*, New York and London: Routledge.

LUFF, P., HINDMARSH, J., HEATH, C. (eds.)
2000 *Workplace Studies: Recovering Work Practice and Informing System Design*, Cambridge: Cambridge UP.

LYNCH, M.
1985 *Art and Artifact in Laboratory Science: A Study of shop work and shop talk in a research laboratory*, London: Routledge & Kegan Paul.

LYNCH, M. & S. WOOLGAR (eds.)
1990 *Representation in Scientific Practice*, Cambridge: MIT Press.

MUCKLI, L., KRIEGESKORTE, N., LANFERMANN, H., ZANELLA, F.E., SINGER, W. GOEBEL, R.
2002 “Apparent Motion: Event-Related Functional Magnetic Resonance Imaging of Perceptual Switches and States”, *Journal of Neuroscience*, 22:RC219:1-5.

MOORSE, M.
1998 *Virtualities: Television, Media Art and Cyberculture*, Bloomington: Indiana University Press.

MURPHY, K.M.
2004 “Imagination as Joint Activity: The Case of Architectural Interaction”, *Mind, Culture, and Activity* 11(4): 267-278.

2005 “Collaborative Imagining: The interactive use of gestures, talk, and graphic representation in architectural practice”, *Semiotica* 156 (1-4): 113-145.

NISHIZAKA, A.
2000 “Seeing what one Sees: Perception, Emotion, and Activity”, *Mind, Culture and Activity* 7(1-2): 105-123.

2003 “Imagination in Action”, *Theory & Psychology* 13(2): 177-207.

OCHS, E., GONZALES, P., and JACOBY, S.
1996 “When I come down I’m in a domain state: Talk, gesture, and graphic representation in the interpretative activity in physicists”, In E. Ochs, E. Schegloff, and S. Thomson (eds.) *Interaction and Grammar*, Cambridge: Cambridge UP.

PEIRCE, C. S.
1934 -1948 *Collected Papers*, Cambridge, MA: Harvard UP.

ROEPSTORFF, A.
2001 “Brains in scanners: an Umwelt of cognitive neuroscience”, *Semiotica* 134: 747-765.

2003 “A double dissociation in twentieth century psychology?”, *Journal of Consciousness Studies* 10: 62-67.

SHNEIDERMAN, B.
1998 *Designing the User Interface*, Reading, Mass: Addison-Wesley.

SUCHMAN, L.
1987 *Plans and Situated Actions: The problem of human-machine communication*, Cambridge: Cambridge UP.

2000 “Embodied Practices of Engineering Work”, *Mind, Culture and Activity* 7(1-2): 4-18.

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

TUFTE, E.R.
1983 *The Visual Display of Quantitative Information*, Cheshire, Conn.: Graphics Press.

TURKLE, S.
1995 *Life on the Screen: Identity in the age of the internet*, New York: Simon and Shuster,

UENO, N.
2000 “Ecologies of Inscriptions: Technologies of Making the Social Organization of Work and the Mass Production of Machine Parts Visible in Collaborative Activity”, *Mind, Culture and Activity* 7(1-2): 59-80.

VOLOŠHINOV, V.N.
1986 *Marxism and the philosophy of language*, Cambridge, Mass.: Harvard UP.

WERTHEIMER, M.
1912 “Experimentelle Studien über das Sehen von Bewegung” (Experimental studies on seeing motion), *Zeitschrift für Psychologie* 61: 161-265.

CHAPTER 6

CONCLUSION: DISTRIBUTED COGNITION, EMBODIMENT AND MULTIMODAL INTERACTION IN MRI PRACTICE

1. DISTRIBUTED COGNITION AND EMBODIMENT

In organizing the theoretical landscape of embodied cognition Margaret Wilson (2002) lists *distributed cognition* as one of the *six different views of embodied cognition*. But, despite its pervasiveness in the literatures on embodied cognition, traditionally the theory has not been primarily concerned with the body. Lindblom & Ziemke (in press), for example, point out that: “The distributed cognition approach proposed by Hutchins (1995) … treats social interactions as directly observable cognitive events, as well as the materials involved in these interactions. However, whereas the body is regarded as one of the media of the information flow, the main focus is not in particular on the role of the body in social interaction, but the transformation of information through different media at a more general level (ibid: 4).”

At the same time, several social scientists have characterized embodied cognition as a view of the body devoid of social aspects (e.g., Goodwin 2003; Sinha & Jensen de Lopez, 2000; Violi, in press). In fact, most views of embodiment in cognitive science do not address the social aspects of cognition and communication. For example, what Zimke (2003) calls (1) *structural coupling* between agent and environment, (2) *historical embodiment* as the result of a history of structural coupling, (3) *physical embodiment*, (4) *‘organismoid’ embodiment*, i.e. organism-like bodily form (e.g., humanoid robots), and (5) *organismic embodiment* of autopoietic, living systems pertain to human, as well as non-human and, in same cases, even non-biological systems. In this respect those views of embodiment are not necessarily concerned with socially and culturally grounded organisms.

When, on the other hand, embodied views do take into account social aspects of cognition (e.g., Barsalou et al. 2003; Lakoff & Johnson 1980, 1999), what primarily concerns them are conceptual structures and higher level knowledge representations realized in the modality-specific areas of the human brain.

Barsalou and colleagues (ibid.), for example, define embodied theories of cognition in the following terms:

> Embodied theories of cognition depart from traditional theories in their assumptions about knowledge representation. In traditional theories, knowledge consists of amodal symbols that *redescribe* sensory, motor, and introspective states … Conversely, embodied theories represent knowledge as partial simulations of sensory, motor, and introspective states … When an event is experienced originally, the underlying sensory, motor, and introspective states are partially stored. Later, when knowledge of the event becomes relevant in memory, language, and thought, these original states are partially simulated. Thus, remembering an event arises from partially simulating the sensory, motor, and introspective states active at the time (pp. 44).

Analogously, to describe the mechanisms of embodiment in social cognition the authors evoke *situated conceptualizations* - “packages of inferences, each tailored to current goals and constraints” (ibid.: 70) - explained in terms of *multimodal simulations*. They claim that multimodal simulations constitute the core of social knowledge inasmuch as “embodied states represent themselves in these constructs” (ibid. 73). For example,

> knowledge about anger resides in simulations of what anger looks like, how one acts, and how one feels introspectively. On this view, simulations of perception, action, and introspection directly constitute the conceptual content of social knowledge. Knowledge is not a redescription of these states in an amodal language, but is the ability to partially re-enact them (73).

Importantly, explanations of social embodiment in terms of internal knowledge representations and their realization in the brain, if not opened to investigations of real-world meaning-making activities, imply: (1) a divide between the body (with the exception of the brain), cognitive (be it embodied or not), and affective states; (2) a reduction of the social world to the stimuli of and their representations in the knowledge system; and (3) a subordination of the body to cognition – i.e. the body is in the function of cognition.

2. MULTIMODAL SEMIOTIC PERFORMANCES IN THE LABORATORY SETTING

To *give the body to distributed cognition* and to *situate the body in a full-blown social setting* the present work focuses attention on the minute instances of everyday practices. It looks at the ways in which we interact with technology and one another to describe how subjective bodies are involved in problem solving, learning, imagining, and how such processes, situated in the world of scientific laboratories, enact and shape such bodies.

The study brings the themes into the spotlight by identifying and describing processes of generation and interpretation of visual representations in MRI. It first shows how visual representations function as fields of cultural negotiations (Chapter 2). The chapter dedicated to the publishing practice points out how the manuscript's authors use MRI images to promote new ways of seeing, generated through the software that they designed. At the same time the chapter discusses how brain images bear traces of standards and expectations of the scientific community at large. In this respect, the images do not just represent the human brain and its cognitive processes; they also function as fields of encounter between local and global, numbers and colours, multiple aesthetics and ways of seeing.

The four chapters that follow continue the investigation of the problem. They, however, address it from the perspective of moment by moment readings of images in the situation of hands-on apprenticeship. Despite the assumption of their transparency, the apprenticeship activities reveal how images become *visible* through processes of coordination between multiple embodied and culturally organized semantic fields. The existence of such semantic complexes reveals the situated character of MRI images – it shows how images are shaped by socio-cultural networks of practices and technologies at the level of semiotic interaction. The chapters also point out how the existence of such complexes allows the practitioners to gain first person experience of the problem at hand. For example, the apprentice can acquire professional knowledge by feeling her body in motion (Chapter 3).

The apprenticeship activity, however, has also to do with laboratory work in that it is not only about "teaching" but also about problem solving. In that respect even the experts' understanding of the digital images often concerns multiple, embodied semantic fields. Practitioners' current knowledge can be grounded in coordination of such fields, or it can be gradually constructed through the involvement with such fields. Practitioners may use their hands to organize representations or to inscribe ephemeral memory traces on them. In other words, multimodal semiotic interaction is not just about "communicating to others". Rather, it shows the interdependence between communication, signification, and inference (Peirce, 1934-48).

Importantly, the involvement of embodied semiotic modalities such as gesture, for example, does not only concern performances of entity-like phenomena - akin to nouns or noun phrases. Gestures can also function as enactments of actions: they narrate the past and envision future events. What's more, they can function as processes of *binding* between multiple semantic fields. Practitioners can represent the brain through a gesture while they perform its "shearing" and the course of its future correction. What is more, gestures can be used to coordinate the brain image on the screen with performances of its transformation, for example. Such processes can be executed through an indexical gesture, an action of touching the digital screen, or an action of physically bringing together multiple visual representations. In such a way, not only individual semantic fields, but also their coordination can be substantially performed in publicly shared physical space (Chapters 3-5).

In fact, the coordination between semantic fields instantiates *conceptual clashes* (Fauconnier & Turner, 2002) and *interpretative chains* (Peirce, ibid.; Eco, 1976, 1984) in the external world of practice. Practitioners coordinate multiple embodied semiotic modalities with the material world of practice to create dynamic and referentially ambiguous performances. Those performances are everyday aspects of practical problem solving in the laboratory setting (Chapter 5).

3. EMBODIMENT, DISTRIBUTED COGNITION, AND MULTIMODAL SEMIOTIC PERFORMANCES IN THE LABORATORY SETTING

Consider multimodal semiotic interaction as part and parcel of laboratory work has theoretical consequences. While it suggests an additional dimension of laboratory studies – the dimension of multimodal semiotic interaction and the body - it also generates alternative ways of thinking about embodiment and distributed cognition.

First, by asking how to talk about the role of the body in naturally occurring processes of meaning making, the study draws attention to the complexities of the subjective body. Inasmuch knowledge structures can be significantly generated in the public space of action, the body reveals itself to be much more than its traces in modality specific brain areas. By showing multiple articulations of the body in respect to its involvement with the culturally rich and socially organized world of practice, the study highlights dependencies among embodied, emotional, cognitive and social realms.

Second, as evident throughout the document, the study is strongly influenced by my background in *interpretative semiotics* (Peirce, 1934-48; Eco, 1976, 1984)[48]. The semiotic perspective makes it possible to note that what practitioners deal with, while they interact with technology and one another, is not simply *symbolic* or *representational* in kind. Rather, it also has to do with imagination, experience and the body.

---

[48] This line of thought has currently been developed in *feminist science and technology studies*

By examining acts of mutual interpretation, interactions and frictions between the semantic fields, the study draws attention to *performances* and *multimodal enactments of semiotic fields*. The performances are ephemeral, dynamic, and often referentially ambiguous or imaginary semantic spaces that arise through practical involvements with the world. They are about the convergence and interdependence between the embodied subjects and the world of practice.

WORK CITED

BARSALOU, L.W., NIEDENTHAL, P.M., BARBEY, A., & RUPPERT, J.
2003 “Social embodiment”, In B. Ross (ed.), *The Psychology of Learning and Motivation,* San Diego: Academic Press, Vol. 43, 43-92.

ECO, U.
1976 *A Theory of Semiotics, Bloomington:* Indiana UP.

1984 *Semiotics and the Philosophy of Language,* Bloomington: Indiana UP.

FAUCONNIER, G. & TURNER, M.
2002 *The Way We Think: Conceptual Blending and the Mind’s Hidden Complexities*, New York: Basic Books.

GOODWIN, C.
2003 “The Semiotic Body in its Environment”, In J. Coupland & R. Gwyn *Discourses of the Body*, New York: Palgrave/Macmillan.

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

LAKOFF, G. & M. JOHNSON
1980 *Metaphors we live by*, Chicago: University of Chicago Press.

1999 *Philosophy in the flesh: The embodied mind and its challenge to western thought*, New York: Basic Books.

LINDBLOM, J. & T. ZIEMKE
In press “The Social Body in Motion: Cognitive Development in Infants and Androids”, *Interaction Studies*.

PEIRCE, C. S.
1934 -1948 *Collected Papers*, Cambridge, MA: Harvard UP.

SINHA, C. & JENSEN de LOPEZ, K.
2000 “Language, culture and the embodiment of spatial cognition”, *Cognitive Linguistics* 11: 17-41.

VIOLI, P.
In press “Beyond the Body: Towards a full embodied semiosis”, In R. Dirven, & R.Frank (eds.), *Body, Language and Mind*, Berlin: Mouton de Gruyter.

WILSON, M.
2002 “Six views of embodied cognition”, *Psychological Bulletin and Review* 9(4), 625-638.

ZIEMKE, T.
2003 “What's that thing called embodiment?”, *Proceedings of the 25th Annual Meeting of the Cognitive Science Society* Lawrence Erlbaum.