Title
Widening the Wideware: An Analysis of Multimodal Interaction in Scientific Practice

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 27(27)

ISSN
1069-7977

Author
Alač, Morana

Publication Date
2005

Peer reviewed

Widening the Wideware: An Analysis of Multimodal Interaction in Scientific Practice

Morana Alač (malac@cogsci.ucsd.edu)
University of California San Diego, Department of Cognitive Science
9500 Gilman Drive MC 0515, La Jolla, Ca 92093-0515

Abstract

This paper analyses multimodal interaction in scientific practice to discuss the possibility of extending the concept of *wideware* to the use of embodied semiotic modalities. It is suggested that the use of gesture, body orientation, gaze, etc. functions as a part of the very process of cognition. The analysis attempts to identify a type of organization that characterizes multimodal interaction among neuroscientists and digital images of the human brain.

Introduction

Contemporary cognitive science, characterized by its research in robotics, artificial life, and dynamical system theory, faces a problem of *higher level* cognition. After largely dismissing explanations of cognition in terms of a physical-symbol system, we still need an adequate mode of accounting for human specific capacities such as abstract and hypothetical thinking, and everyday meaning-making in culturally rich environments. Andy Clark's (1998; 2001) solution to the problem is *cognitive technology* or *wideware*:

> The central idea is that mindfulness, or rather the special kind of mindfulness associated with the distinctive, top-level achievements of the human species, arises at the productive collision points of multiple factors and forces – some bodily, some neural, some technological, and some social and cultural. As a result, the project of understanding what is distinctive about human thought and reason may depend on a much broader focus than that to which cognitive science has become most accustomed, one that includes not just body, brain, and the natural world, but the props and aids (pens, papers, PCs, institutions) in which our biological brains learn, mature, and operate (Clark, 2001: 141).

Like Daniel Dennett (1995, 1996), Edwin Hutchins (1995), and Kirsh & Maglio (1994), Clark emphasises the crucial importance of the environment, particularly language and culture, for human cognition.

The present work further explores the concept of cognitive technology. It provides additional empirical evidence for existing theoretical claims, and focuses on possible extensions of the concept to a variety of embodied semiotic modalities such as gesture, body orientation, gaze, etc. It suggests that these modalities not simply serve to express our internal cognitive processes, but also to develop and perform such processes through public actions.

Because of its intrinsic socio-cultural character, in addition to careful experiments (e.g., Kirsh & Maglio, *ibid.*), cognitive technology needs to be studied by examining everyday practices. The present work does so by combining methods of *cognitive ethnography* (Hutchins, *ibid.*) with micro-analysis of multimodal interaction (Goodwin, 1994). The micro-analysis is applied to data captured with digital video. The video data allow for a monitoring of activity in rich environments of practice. Analytic judgments concerning analysis of a minute instance of practice are warranted by findings from long-term ethnography.

Because of complexities inherent in the analysis of everyday practices, it is legitimate to wonder if such an analysis could permit an identification of patterns in human action. In other words, is it possible to detect systematic organization among multimodal elements of micro-interaction?

Analysis

To tackle these questions, I carried out an ethnographic study of scientific practices in three laboratories of cognitive neuroscience. All three laboratories use Functional Magnetic Resonance Imaging (fMRI) technology as a principal method of investigating the human brain.

According to Kutas *et al.* (manuscript), it is commonly believed that the appeal of fMRI technology is influenced by its apparent transparency, producing an impression of reduced need for interpretation. Despite this impression, visual representations of experimental data are always in need of further interpretation. Analysis of the details of local practices suggests that brain images do not function as self-explanatory representations that simply support scientific reasoning. Knowledge acquisition and the comprehension of experimental data are accomplished through complex interactions of cognitive, bodily, and socio-material means (Alač, 2003; 2005; Alač & Hutchins, 2004).

To investigate how these means coordinate through practice, I explore the idea of *multimodal interactional systems* (MISs). MISs combine various semiotic modalities, e.g., speech, gesture, body orientation, gaze, and visuo-spatial representations, to produce meaning. They are constructed on-line in the public space of action through interaction among multiple social actors and their environment of practice. I investigate a kind of MIS that allows visually represented digital information to become an embodied social experience.

Figure 1 is a representation of such an MIS. The system is a meaning-making cycle that can be roughly divided into three phases. During the first phase the interlocutors establish shared attention toward the digital image. This phase initializes the *translation* of the 2-D world of the digital screen into the 3-D world of action. Note that the digital image itself is a part of the cycle. The immersion of the image in the cycle points out that the meaningfulness of the image does not exist when the image is taken in isolation. The second phase is denominated “performative/embodied action”. Performative/embodied action creates *enacted processes,* or, alternatively, appropriates pre-existing, static representations. An enacted process translates visual, 2-D representations, into embodied performances expressed through gesture or body movement. The performance functions as an interpretation, or a temporary solution to the problem that the visuo-spatial representation presents. Gestures or body movements that participate in the performance do not simply generate a site of reference. They add new layers of meaning to visual representations, for example, three-dimensionality and mobility. During the third phase of the cycle, selected and interpreted aspects of the image are coordinated with the digital screen. When this stage is completed the interlocutors can understand what is present on the screen. The cycle generates *visibility* through the three stages. All three stages are tightly coordinated and mutually dependent on each other.

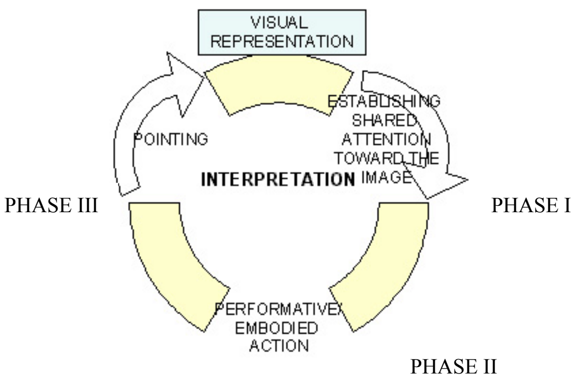

Figure 1

To illustrate how such MIS is instantiated in a real-world practice, rather than focusing on a single example[1], I analyze and compare three instances of apprenticeship interaction. The instances are representative of a larger corpus of data (9 hours of video tape). They were video-taped in three different fMRI laboratories across 12 months of ethnographic study.

Example 1: Retinotopy

The first example concerns an activity where the novice acquires the capacity to identify retinotopically organized areas on the digital brain representation. The participants view on the computer screen what is called a “phase map”, i.e., the digital representation of the neuronal activity in the visual cortex (for a more comprehensive description of the setting see Alač & Hutchins, 2004).

Here is a brief excerpt from the interaction in which the expert explains to the novice how to identify the first visual area (V1) on the map[2]:

1 So probably this is the center [touches the screen with her index finger] (0.1)
2 right here [takes the sheet of paper with the drawings and moves it clockwise and points to what represents the fovea on the chart]
3 And when we look at this map it looks something like that [picks up the paper and holds it next to the computer screen] (0.5)
4 So V1 is gonna be in the center [briefly traces the borders of the V1 representation on the chart with her index and middle finger]
5 it’s gonna be this pie shape it’s probably covering approximately this area [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] (0.5)
6 Ok?

Before the interaction took place, the expert drew a chart exemplifying the mapping between the visual field and its projection onto the cortex. Even though the chart was manufactured by the expert herself, it functions as a collective depository of knowledge. Because the chart represents shared knowledge of the research community, and because its structure was already explained to the novice, the chart functions as a *known* structure. This known structure must be mapped onto the brain image that is still an *unknown* territory.

The mapping is accomplished through the formation of an MIS. In line 1, 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. Counterintuitively, she first points toward the unknown (i.e., the digital image) and only then toward the known structure (i.e., the chart). This pointing sequence suggests that the pointing toward the computer screen is a general process of attention directing, rather than an act of referring to a particular structure on the image. The pointing forms the first phase of the MIS (Figure 2).

Importantly, this pointing also functions as a preparatory stage for the second phase of the process. To make the mapping between the chart and the brain image more effective, the expert rotates the chart (line 2). The rearrangement of the environment allows her to prepare for a more straightforward alignment between the chart and the image, making memory and recognition a simpler, visual task.

The role of the chart and the mapping of various representational forms are further developed in lines 3-5. In line 3 the expert picks up the chart 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 novice to see the structures present in the two representations as comparable. Next, a particular element of the chart - the representation of the visual area V1 - is *mapped* from the chart onto the brain image. A crucial part of the mapping process is played out by gesture. In line 4 the expert places her right hand onto the chart and briefly traces with her index and middle finger the borders of V1. The coupling of the gesture with the chart and the speech *enacts* V1 in the environment of practice (phase two in the MIS). The gesture attributes a dynamic form to static visual representations, and makes its specific elements particularly salient.

Finally, in line 5 the hand is lifted from the chart, moved toward the brain image and carefully placed onto the

[1] For such an analysis, see Alač, 2003; Alač & Hutchins, 2004, Alač, 2005.
[2] Transcription conventions follow Goodwin (1994).

image. The gesture and its movement from one representation to the other function as a mapping process that binds the two physically separate representations together. Once the hand is placed onto the image, another gestural enactment of V1 is performed. The gestural enactment inscribes dynamic, ephemeral marks on the image. In this way the gestural form functions as an interpretation that selects, and makes specific elements of the brain image salient. While enacting the V1 representation on the brain image the expert utters: "it's gonna be this pie shape is probably covering this area." The categorization, performed not only through speech, but also via gesture-image coordination makes the MIS complete.

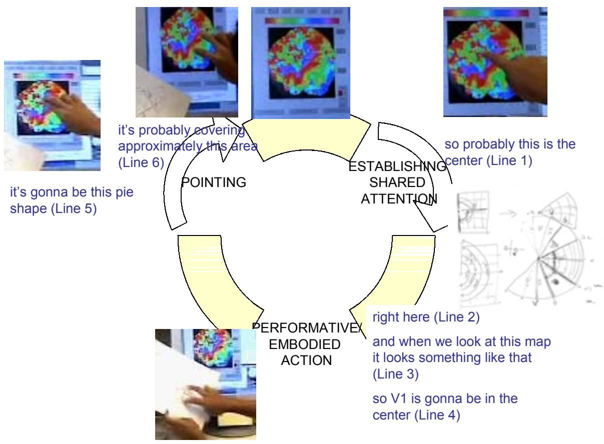

Figure 2

Example 2: Shearing Correction

This example examines an interaction between the principal investigator (PI) and the post-doctoral student (PD). The excerpt was recorded during an early stage of fMRI 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.

By using the computer mouse the PI is able to switch between functional and structural images. This rapid switching between the images is done in order to make apparent potential differences between the two. The PI notices a "shearing" in the functional image. 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."

Before the researchers are able to correct the image distortion, they have to identify and characterize it precisely. The following excerpt illustrates the identification and characterization of the image distortion:

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: you can see=
4 it's sort of [gesture exemplifying distortion] ((PI looks toward PD))
5 PD (affirms, but what she said exactly cannot be identified)
6 PI It's /sort of=
7 PD /Yeah
8 PI it's actually sheared=((PI's left hand is still in the position assumed in the previous gesture))
9 it's going up this way [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]
10PD Right.

In the beginning of the excerpt (line 3), the PI directs his interlocutor's attention towards the computer screen by uttering "you can see." The utterance announces the centrality of the digital image for the constitution of action that follows. The PI enhances this shared tuning of attention towards the computer screen by leaning in this direction. The PD can read the PI's body orientation as an invitation to direct her gaze toward the same site. Through this tuning of attention toward the computer screen the first phase of the MIS (Figure 3) is accomplished.

Next, the enactment of the brain image distortion takes place (i.e., phase two of the MIS). The cognitive task for the actors is to see the distortion in the image. Yet, it is quite difficult (especially to a non-expert eye) to identify the distortion if the image is taken in isolation. The visibility of the distortion is enhanced through the interaction of the image with speech and gesture. In line 4, the image is characterized as distorted. The characterization is produced through coordination of the linguistic expression "it's sort of", with the computer screen, and the gesture. The linguistic expression provides general frame for the action, while the digital display and the gesture coordinate and designate the distortion through its performance. The gesture *chooses* particular elements of importance in the display. It makes salient only those features that are relevant to the present task, i.e. its distortion. Moreover, the gesture not only indicates that something that is 3D and has a round shape is distorted, but it also illustrates the way in which it is distorted: its left half is positioned higher than the right one. It is obvious that the researchers know that the "shearing" of the image is a consequence of an abstract computational transformation. Even so, their gesture indicates that they tend to think of the problem in terms of concrete, physical action. In other words, the digital brain image, coupled with gesture, enable scientists to accomplish the brain-mapping tasks in terms of an embodied action that takes place at a level comparable to physical, real-world engagement (Fauconnier & Turner 2002; Lakoff 1987).

Through lines 5-9 the two participants further coordinate their knowledge: the PI makes sure that the PD saw what he saw. Notice that during the activity that takes place in lines 5-9, the PI's hand remains in the position assumed during the previous gesture. This steady position of the hand is important. It allows for the linking of the enactment carried out in line 4 with the subsequent performance, through which the PI indicates the nature of the distortion directly on the image. This linkage is further developed through the PI's pointing gesture. When the PI places his left hand directly onto the computer screen, he briefly points with his index finger to a portion of the image. Through this gesture, the now visible, performed distortion is *transposed* back onto the digital representation. The loop from the computer screen to the distortion enactment and back to the computer screen is now closed.

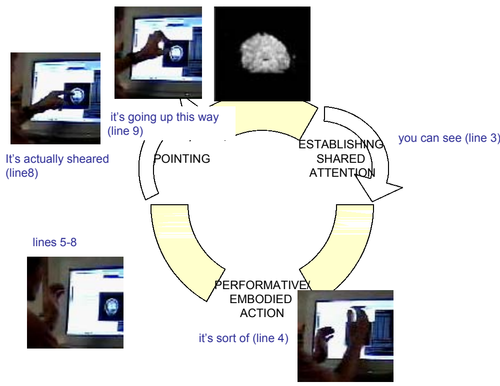

Figure 3

Example 3: Motion Correction

The third example describes once again an interaction where the expert (E) explains to the novice (N) how to asses the existence of artifacts in the experimental data (for a more comprehensive description of the setting see Alač, 2005). As the excerpt from the interaction will attest, here the nonalignment between the images is explained in terms of the movement of the experimental subject that caused it: the artifact (i.e., *motion artifact)* is caused by subject's movement.

1	E:	That's definite I can see her in this plane= [Points to the computer screen]
2		going from here to here. [Sweeps her arm in a downward motion and halts at a certain point]
3		Aaaa ((disapproves))
4	N:	(Aaaa; this is good one) [Points to the sagittal view of the brain slice on the screen]
5	E:	Slice, edit ((instruction for pressing button))
6	N:	(Oh, here we are)
7	E:	Is she moving any more in 30? ((Moves closer to the screen))
8	N:	She's going - = ((Hunches))
9		Like this one [Points to the axial view of the brain slice on the screen]
10		Is going down. Hhhhh ((laughs)) [Swings with right hand palm down downward]

The excerpt illustrates how the expert guides the novice in spotting the nonaligned representations by associating them with movements that the subject made in the scanner. This evocation of the movement can be easily detected in the participants' speech. The expert verbally explains the nonalignment of the images by using linguistic forms that express motion. In lines 2 the verb "to go" is used, while in line 7 the verb "to move" is used. These expressions of movement are in clear contrast with the static character of the images.

But the movement of the subject is not only evoked through speech. By manipulating the computer screen, i.e., changing the display of a particular brain slice through time, the movement of the subject becomes *visible*: "I can see her in this plane", line 1. With the difference between two brain images, the participants are able to perceive the apparent motion.

The full conception of the movement, however, is only recreated through its performance in the 3-D world of action. In line 2, the expert quickly points toward the computer screen and consequently transforms that indexical gesture into the gesture of the subject's movement. Particularly interesting is the capacity of the gesture to perform the hypothetical action. Note that what is believed to be the cause of the nonalignment (the movement) can only be inferred from its effect (the nonalignment of images). The movement of the subject could not have been seen or experienced by the researchers while the subject was lying 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 researchers supervise what goes on in the scanner. Yet, by performing the subject's movement through the gesturing hand, the previously unseen behavior of the subject is recreated in the physical space shared by the participants.

This instantiation of the movement becomes even more clearly visible in lines 7-10. Because of the expert's movement toward the computer screen (line 7), one can infer that she cannot easily see what is happening on the screen. At the same time her body movement functions as an attention-channeling device. The movement and its coupling with speech and the computer display instantiates the first phase of the MIS (Figure 4).

In response to the expert's question, the novice takes the floor and actively engages in the creation of the subject's movement (lines 8-10) (phase two of the MIS). Through the enactment of the movement the participants tackle the problem appearing on the digital display.

Interestingly, in line 8 the novice's body assumes a primary role in the meaning-making process; it becomes a 3-D performative sign of the subject's behavior in the scanner. The fact that in hunching, the novice's body functions as a model of subject movement is confirmed at the linguistic level. The expression "She is going" parallels the body movement and expresses the idea of the subject's body in motion.

In this way, rather than dealing with an apparent motion, the novice creates a process where her body moves as if it was the subject's body. The nonalignment becomes understood by experiencing, first person, the subject's hypothetical movement. The participants can not only perceive the movement caused by the manipulation of the computer screen; they can reason about such a movement by experiencing it through the movement of their own body. In this sense, the novice's body movement is used not only to communicate, but also to make sense of the problem to be solved.

Next, the enactment of the moving body has to be linked back to the images. This is achieved through a pointing gesture. In line 9, through the indexical gesture, the hunching combines with the brain image on the computer screen (phase three of the MIS).

However, the cycle is not yet fully closed. In line 10 the indexical gesture is quickly transformed into another hand gesture of the subject's movement. Now the novice's arm is mapped onto the subject's head. This second performance through the hand gesture is not superfluous. While the hunching gesture has the advantage of preserving the straightforward mapping between elements of two domains (i.e., the novice's body parts map onto the subject's body parts: shoulders map onto shoulders, head maps onto head, etc.), the hand gesture can more precisely signify the direction of movement that the subject performed in the scanner. In addition, because of its orientation and shape, the hand gesture contains some elements of indexicality that the hunching movement lacks. The hand gesture, while it stands for the subject movement, also directs participants' attention back to the computer screen.

What is remarkable about this 3 seconds long activity is its dynamicity. The novice's body, her arm (as an indexical sign and as a performance of the subject movement), her linguistic utterance, as well as the visual representations on the computer screen, are combined in the construction of the motion artifact. None of these forms is an all-encompassing signifier. For example, the subject's movement is represented: 1) by the apparent motion created through the difference between images; 2) by the novice's body; 3) through the novice's gesture. Each form is a partial view that highlights different aspects of the subject's movement that has to be seen on the computer screen. Importantly, the passage from one form to the other is not a linear, unidirectional path from problem to solution. Rather, the participants build understanding through bi-directional coordination across different elements of the cycle. The result of such coordination is an acquisition of new perspectives and enhanced understanding.

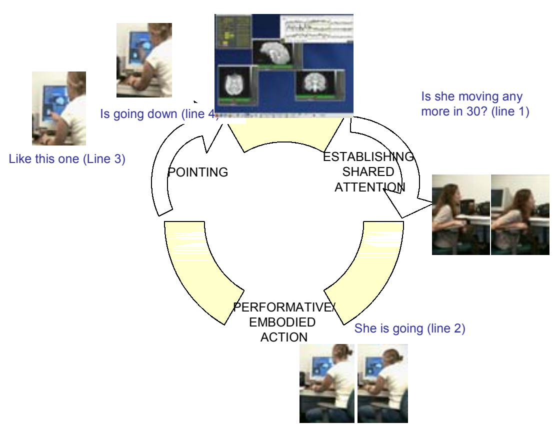

Figure 4

Discussion

It is well known that human behavior is highly organized. Nevertheless, the exact details of such organization are still unknown. By paying careful attention to details of on-line, multimodal interaction we can start to identify the organization of such interactions. To do so we identify their properties (often dynamic and flexible), as well as ask how such properties are coordinated and how they enable and facilitate each other. This can be accomplished only by understanding an interaction as embedded in a real-world practice. In this paper I identified an organization of micro-interaction, where the acts of *reading* 2-D digital images involve embodied performances. This type of organization allows information to move from the computer screen to the space of social embodied action, and back to the computer screen (Figure 1). The movement across the multiple meaning spaces and their co-articulation generates understanding.

The analysis of the three examples shows that MISs that share the same general form can be rather diverse. They vary in respect to the types of semiotic elements involved in the process. The first stage of the cycle may, for example, be accomplished through the coordination of the digital display with an indexical gesture, linguistic expression, or body movement. In Example 1, the expert directs the novice's attention by pointing toward the screen. In Example 2 the same function is performed via the linguistic expression, "you can see." In Example 3 the expert moves her body toward the screen. She does so to enhance her own seeing, while engaging the novice in the shared tuning of attention.

The diversity and complexity of the interpretation cycle, however, is most clearly manifest in the second, performance phase. The elements involved in the accomplishment of the phase range from relatively static to highly dynamic ones. Example 1 shows how stable representations, such as the expert's chart, can function as central carriers of meaning production. By involving the chart in the process, the cycle adopts knowledge structures of an entire scientific community. The stable representation is, at the same time, rendered dynamic through its coordination with the gestural actions. In Example 2, on the other hand, the enactment is primarily performed through a gestural form. The gesture, in coordination with other semiotic modalities, provides an embodied account of the abstract experimental data. Such an account allows scientists to think about the problem to be solved in terms of a direct manipulation of a 3-D object. Similarly, the enactment phase of Example 3 is highly dynamical. The enactment engages the body of the junior practitioner in the construction of understanding.

Also the means used to accomplish the third phase of the cycle are quite diverse. In the third phase of all three examples the indexical gestures were directed toward the digital screen. Rather than simply pointing toward a location on the screen, the indexical gestures function as a mapping device. They coordinate what was generated through the enactment stage with the digital brain images. Once the cycle is completed, the participants can carry out various enactments directly on the computer screen. The gestural enactments inscribe onto the image the structure to be seen by the practitioners. It is essential to notice that to grasp the global meaning of the process, rather than focusing on one semiotic modality (as is frequently the case in *gesture studies*; for review see McNeill, 2002), it is necessary to trace ways of co-articulation among multiple modalities.

Conclusion

Interest in visual representations as structures that stabilize knowledge is not new in cognitive science. Since the early works in *diagrammatic reasoning,* we have known that the way in which a problem is represented – whether it explicitly preserves the information about the topological and geometric relations among the components of the problem - is crucial to our performance in problem-solving (Larkin & Simon, 1987).

Yet, we still don't know how such processes actually take place through everyday social practice. A close analysis of video-taped scientific practice confirms that digital displays are powerful cognitive tools: scientists reason about an enormous amount of numerical, abstract data that are re-presented into a form that exploits the specific computational powers of the human visual system (Kirsh, 1992; 1995; Clark & Thornton, 1997). However, such an analysis also points out that digital displays do not function in isolation. Meaning emerges from an activity where not only eyes, but hands and bodies (Latour, 1986), are actively involved.

Observations of such activities and their involvement in cognitive tasks suggest that the idea of wideware could be extended to embodied semiotic modalities. For instance, the gesture, its persistence through time, and the active use of space, participate in processes of mapping among various external representations (e.g., Example 1). The problem of mapping or binding disparate representations into units is one of the central questions in cognitive science (for discussion see Fauconnier & Turner, 2002). Here we see such cognitive process being largely executed in the external and shared world. The examples also show how gesture, in coordination with visuo-spatial representations, plays a crucial role in processes of selection. Gesture generates transient inscriptions on the rich visual representations (e.g., the third phase of all three examples) to articulate their specific features relevant for the task at hand (e.g., Goodwin, 1994). Such inscriptions create categories, organize the environment into meaningful units, and thus transform the complex internal task into a simpler, socially shared, embodied process. What is more, gesture and body movements can be used to enact imaginary and hypothetical events in the environment of social practice (e.g., shearing as a physical process in Example 1, and the previously not-seen subject's movement in Example 3). In this way the junior practitioner, who cannot just by looking at the computer screen infer the position and form of relevant structures (e.g., the V1 borders in Example 1, or the non-alignment among images in Examples 2 and 3), can observe and experience the hypothetical cause of the problem, and its solution, as well as the process of its production in the public environment of practice. One of the processes that allow him/her to do so is the coordination of his/her body movements with the expert's action and the computer screen (e.g., Example 3). Moreover, as all three examples indicate, he/she can reason, and participate in scientific practice by engaging the embodied actions of social partners. Such actions become a part of his/her wideware.

Acknowledgements

I would like to thank the following people for their contribution to this project: Ed Hutchins, Wendy Ark, Geoff Boynton, Lisa Cartwright, Aaron Cicourel, Seana Coulson, Hollie Crower, Ryan Downey, Jeff Elman, Yrjo Engestrom, Gilles Fauconnier, Deborah Forster, Charles Goodwin, Ed Hubbard, Maurizio Marchetti, Ayse Saygin, Marty Sereno, Betty Shor, Joan Stiles, Hsin-Hao Yu, and the participants in the ethnographic study.

References

Alač, M. (2003).Squashing, Rotating, Seeing, and Going: On visual knowledge in fMRI research, Proceedings of the 25th Annual Meeting of the Cognitive Science Society.

(2005). From trash to treasure: learning about the brain images through multimodality, *Semiotica*, 156-1/4.

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

Clark, A. (1998). Where brain, body, and world collide. *Daedalus,* 127(2), 257-280.

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

Clark, A. & Thorton, C. (1997). Trading Spaces: Computation, Representation, and the Limits of Uninformed Learning. *Behavioral and Brain Sciences* 20, 57-90.

Dennett, D. (1996). *Kinds of minds*. New York: Basic Books.

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

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

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

Kirsh, D.(1995). The intelligent use of space. *Artificial Intelligence* 73, 31-68.

Kirsh, D. & Maglio, P. (1994). On Distinguishing Epistemic from Pragmatic Actions. *Cognitive Science*, 18, 513-549.

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.

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

Latour, B. (1986). Visualization and cognition: Thinking with eyes and hands, *Knowledge and Society*, 6: 1-40.

McNeill, D. (ed.) (2002). *Language and Gesture*, Cambridge: Cambridge UP.