Squashing, Rotating, Seeing, and Going:
On Visual Knowledge in fMRI Research

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

Abstract

In order to describe how visual representations are used in brain mapping research, two concrete instances of fMRI practice are analyzed. The primarily focus is on the correlation of distributed representational media that build conceptual structure of the participants in action. The analysis shows that comprehension is organized through complex conceptual structures that transform abstract concepts into manipulable, object-like entities. It is also observed that such structures employ phenomena such as fictive motion in which a static entity is construed as being in motion, as well as hybrid, non-veridical conceptual construct.

Introduction

In cognitive neuroscience, functional MRI is used to visualize brain functions by visualizing the local changes in magnetic field properties occurring in the brain as result of changes in blood oxygenation. This technique provides both images of brain structure, and "maps of active processes within the brain," i.e. maps of brain function.

The question is, how are these images used during the scientific practice? In other words, how do the visual representations of brain function become understandable and meaningful for scientists? It is often argued that the appeal of fMRI technique is influenced by its apparent transparency, producing an impression of reduced need for interpretation. The obvious question then is, does the meaningfulness of fMRI brain representations reside in the image itself?

Related to these questions is the problem of characterizing differences in working with brain images versus directly handling "real brains." Are scientific practices that deal with digital representations radically different from those that directly handle physical tissues and organs? Is interacting with digital images an abstract procedure, or does it enable scientists to accomplish brain mapping tasks in terms of embodied action that takes place at a level comparable to physical, real-world engagement?

Anne Beaulieu (2002), in her work on human brain mapping research (involving fMRI as well as PET techniques) highlights the spatiality of PET/fMRI measurement and importantly notices that analysis that brain mappers conduct relies on spatial components and anatomical referents. However, the point of her work is to show that a paradox lies at the basis of human brain mapping research. The paradox is that the very same cognitive neuro-scientists who rely on visual representations in their work, reject the importance of these representations. According to the researcher that Beaulieu inteviewed, brain images have a "popularizing and iconic role," that is, they are used in the promoting "attractiveness" of the research on the brain as well as simply "adorning" publications concerning psychology and neuroscience. However, their direct involvement and importance in the research is dubious, at least from the point of view of the researchers:

> Given that a wealth of visual representations accompanies this new approach to the study of mental phenomena, what is the role of images in these developments? What does it mean that mental functions become visible? The answer given to both questions by researchers involved in brain mapping would be this: very little. They would generally argue for a highly circumscribed role of images in brain mapping; brain mappers insist that they do not know the brain through images and that these "pretty pictures" are at best useful visual aids when giving talks (Beaulieu, 2002: 54).

The attribution of "highly circumscribed role to images" and their characterization as "useful when giving talks," suggests an implicit belief in a dichotomy between communication and thinking. In the present work, I will argue that this dualism between external representations considered as optional and a posteriori, and internal representations viewed as fundamental and exclusive elements of cognition that takes place in "minds" construed as somehow separate from or outside of the world, is a false dichotomy. The inadequacy of this presupposition is revealed when analysis is applied to effective, real-world practice of fMRI brain mapping.

My analysis does not attempt to prove or disprove the importance of brain images as visual data, or to explain why such paradox may exist. Instead, by documenting the fMRI brain mapping activities, that is, by showing how the visual images are used in the actual practice of fMRI brain mapping, I described the way in which the visual form of

such data gets actively involved in the conceptualization and production of meaning.

Method

The following analysis is part of a broader ethnographic study of the production of visual fMRI evidence. The study involves observations of scientific practices conducted in two laboratories. Both laboratories' primary research interest is in the visual system. The study took place over a period of 6 months, and includes a variety of data collection methods including direct observation, video, semi-structured interviews, and analysis of documents such as scientific papers, laboratory manuals, and scientific correspondence.

Analysis

The instances analyzed here are examples of general patterns observed across actors and laboratories.

Seeing Colors to Identify Borders

The following excerpt concerns charting retinotopic maps in the visual cortex. I analyze the interaction between an expert (E) and a novice (N), where the novice is being taught how to analyze fMRI retinotopy data, that is, to "see" specific brain areas in a digital representation of the neuronal response to visual stimuli.

Retinotopic mapping is a procedure for detecting the ways that the brain maps a visual scene. One can speak of retinotopic map when adjacent locations in the visual field are represented in adjacent locations in the cortex. The process of retinotopic mapping 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 stimuli and its neuronal response, scientists are able to identify which part of the visual cortex processes stimuli located at specific points in the visual scene. This allows scientist to assess the location and borders of a specific brain area, i.e. attribute meaning to the brain image. The representation of time series of neuronal responses is color coded on the map (Fig. 3). Hence, in order to identify the borders of visual brain areas, the researchers have to identify a change in color.
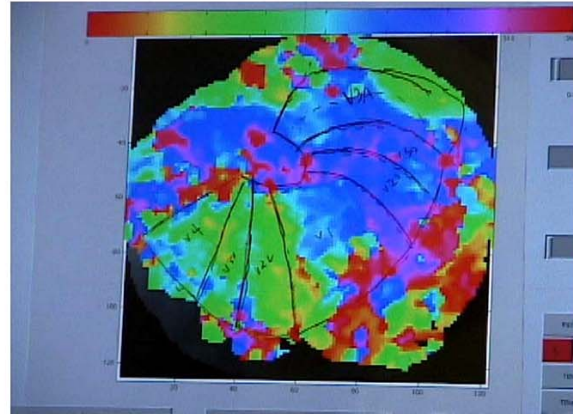

Figure 3: Activation map with inscribed borders of visual areas.

The dialogue below is an excerpt from the activity of looking at the *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 about fixation. Ideally, this type of stimulus provokes the alternation frequency of the neuronal response that is the same for all points in the visual field, but varies in the temporal phase of the response.

5 - 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]
6 - N: Mhm.
7 - E: Mmmm [hesitates] then it goes out to purple and then back to orange and
8 - then out to purple again.
[still pointing and tracing the lines]
9 - N: I Actually I can see that now. Hhhh
10 - E: You kind of see some of that intermediate (stuff where) it goes from
11 - orange to red to purple, right?
[points on the computer screen]
[the novice is talking at the same time, but his utterance is incomprehensible]
12 - Yeah. Ok. Right
[the expert takes her hand away from the computer screen, claps over the table and quietly laughs]
13 - If you (would) believe me (it would be) very nice. Hahaha
14 - N: No, I see it. I see what you are saying, I see what you are saying.
15 - E: Haha. Ok.

This passage is a clear example of the important role the visual form of fMRI data plays in understaniding configuration of retinotopic maps in visual cortex. The capacity to see the transformation of color and the reversal in the direction of this transformation is fundamental for the identification of borders and separation of visual areas.

The act of teaching the novice to see retinotopic maps is performed through the coordination of linguistic, as well as what are at the same tie indexical and iconic gestural representations. For example, the expert indexes and traces the borders of the visual areas (lines 5 and 8). Bodily movements and expressions of emotions, such as the expert clapping over the table to signal jokingly her determination, or laughing (lines 12 and 15), are also involved.

Even though the colors on the phase map do not move, i.e., the phase map is factively a static representation, the subjects conceptualize the map fictively in terms of movement. In fictive motion (Talmy, 2000) one deals with nonveridical phenomenon 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. This is seen clearly in lines 7-8 and 10-11 where the expert speaks of *one color is going into the other,* even though it is clear that the colors are not moving.

As Talmy has pointed out, the phenomenon of 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 10 the expert states: "you kind of see some of that intermediate (stuff where) it goes from ... ." While moving their gaze over the map, the participants act as if they see the progressive change in the static scenery. The fact that the actors can make sense of the experimental data in terms of fictive motion of colors allow them to process the data in a way that is particularly suitable for human cognition. If the same data were represented quantitatively, it would be very hard, or potentially impossible, to determine the propagation of what scientists call "the traveling wave of neuronal activity".

Identification of meaningful entities such as visual area borders through fictive motion of colors indicates that these scientists' conceptualization depends on the visual form of fMRI data presentation. The image is not only a "useful visual aid when giving talks," but a crucial component of conceptual structure. The representation of experimental data in visual form allows for more apt conceptual processing of such data that include in its structure not only veridical but non-veridical phenomena as well.

In addition, the passage shows that the identification of such areas is not as immediate as one might expect. The expert has to make sure that the novice is able to see this fictive movement. In line 9 the novice assures her that he can see it: "I Actually I can see that now." The expert insists again in highlighting what the novice should see.

In lines 12-13, 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 *voices* the laboratory as well as the larger neuroscience community, believing what she says situates seeing in the enormously complex structures that construe it.

The novice in his reply (line 14): "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. 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 not somehow direct, but is the product of the expert's *saying.* And the expert's *saying* involves a coordination of multiple structures built through the situated human interaction of learning to see.

The novice's expression can be interpreted as representative of the idea that underlies the fMRI practice in general. He not only says that he understands what is being said, but that he understands through seeing, and hence his expression points out the close mapping between the domain of understanding/thinking, and seeing. The fMRI technique, by spatially identifying areas where particular brain processes take place, allows researchers to infer the nature of specific cognitive processes, and hence can be described as based on the metaphor of "seeing is believing". However, this does not mean that such technique is based on a reduced need for interpretation. This analysis shows the meanifulness of fMRI brain representations does not reside in an image itself, but involves complex network where different representations are coordinated. Such representational networks comprise veridical as well as non-veridical elements that allow for a more apt understanding of scientific phenomena.

Manipulating Images or Physical Objects?
The second example is taken from an initial stage of fMRI data analysis. A senior fMRI researcher (R) and two advanced graduate students (S) are seated in front of one of the laboratory computers. In the excerpt below, a small functional image is displayed on the center of the screen and the researcher is explaining to the students how to correct the distortion in the image. Because of an inadequacy in the software program, the image series cannot be automatically corrected; it has to go through a subsequent transformation. Modifications are digitally performed on the image in four stages (here denominated - rotation, squashing, un-rotation, un-squashing).

1 - R: 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]
2 - and then squash
[places both hands on the screen as if holding the object represented and slightly pushing it from both sides]
3 - and then un-rotate
(rotates both hands holding the imaginary object towards the left)

4 - and un-squash
[quickly moves right hand as if mimicking the expansion of the representation]
5 – S: Hmmm.
6 – R: That's about … you may want to do 40 degrees. (It's about 45 degrees)[1]

In the passages the PI speaks and gestures as if dealing with a physical object that can be directly manipulated. He speaks of "rotating" and "squashing" in reference to the two dimensional digital image, which obviously cannot be directly and physically rotated or squeezed. The researcher reaches toward the computer screen and acts as if holding a round object and moving it towards left/right (Figure 1). Notice that the gesture specifies the manner of the action, not expressed by the linguistic form. The gesture and the linguistic expression produce complementary, rather than redundant information. The researcher also performs the action of "squashing" over the image, as if squeezing a three-dimensional object (Figure 2).

However, the transformation is possible only through computational means that produce an image which appears as if it were physically manipulated. The actual action, i.e., mathematical processes computed and buried in the machine, generates a new image representing the brain in a new position (e.g., a horizontal, rather than vertical position as if it was physically rotated). This appearance of the new position enhances the idea that the image was manipulated physically, even though its transformation has been achieved through abstract means.

Figure 1: "And then un-rotate."

Figure 2: "And then squash."

But how is this possible? What is the conceptualization behind such a process? What the researchers have in front of them are two dimensional static objects. The only element in motion is the hand touching the computer screen. However, the researchers talk and gesture, and probably perceive the image whose distortion has to be corrected as something that can be rotated or squashed. In order to conceptualize the image as such, the researchers employ a process known as *conceptual compression* (Fauconnier & Turner, 2002).

Conceptual compression derives from the theory of *conceptual integration*. According to Fauconnier and Turner (1998), conceptual integration operates over mental spaces as inputs and makes use of a four-space model. Mental spaces are short-term mental scenarios constructed on-line as a sub-structure of a given domain. The conceptual integration networks include two input spaces plus a generic space, representing conceptual structure that is shared by both inputs, and the blended space, where material from the inputs combines and interacts. In blending, structure from two input spaces is projected to a separate space, the "blend". The blend inherits partial structure from the input spaces, and has emergent structure of its own. Blending is an on-line, real-time process that creates new meaning through the juxtaposition of familiar material (for review see Coulson & Oakley, 2000).

The knowledge of the context of action, as well as information on the usage of gestures and language, suggest that in the activity of distortion correction, the mechanism of conceptual integration allows the researcher to conceive of two different conceptual entities as if they were an integrated item. In this way, the physical brain and its visual representation on the computer screen are manipulated, as if they were a single item. The link between the digital brain slice representation and the physical brain is representational and metonymic. The digital brain slice represents the physical brain slice, and the physical brain slice is in metonymic relationship (part-for-whole) with the whole physical brain. The integration of two entities, the digital representation and the physical

brain, produces an imaginary conceptual structure (denominated "blended space" by Fauconnier & Turner, 1998) that allows the co-existence of an impossible combination of characteristics - the imaginary structure 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.

In order to construct this imaginary conceptual structure, where one is communicating about and manipulating an entity that is at the same time two- and three-dimensional, digital and physical, the digital brain slice image is crucial. The image assumes the role of a *material anchor for conceptual blend* (Hutchins, 2002). Hutchins analyses conceptual integration where one input is a conceptual space in the usual sense (here the real brain, i.e., the referent of the brain slice image), while the other one has the structure of a physical object or event (here, the brain slice image)[2].

However, it is still unclear why the actors gesture toward the image on the computer screen, as if holding and manipulating a round object, if they are dealing with a conceptual, imaginary entity? I believe that the answer to this question lies in the capacity of the gestural expression to externalize the imaginary conceptual structure[3]. It can be argued that, by looking at the PI's gesture, we "see" an external conceptual structure that contains the impossible combination of characteristics. The gesture communicates and manipulates an entity that is at the same time two- and three-dimensional. The presence of such a conceptual construct can be inferred from the gestural form that manipulates something appearing on the computer screen, while, at the same time, allowing for such an entity to be physically rotated and squished. The gesture deals not only with the image, since it performs the action of three dimensional physical manipulation. However, its form is shaped by the appearance of the image itself[4].

At the same time it should not be forgotten that the gestural expression accomplishes the externalization of the imaginary conceptual structure only inasmuch as it is coupled with other representational forms that partake in the action. On the one hand, the PI's gestures are accompanied by linguistic expressions (i.e., "If you rotate 30 degrees this way", line 33; or "and then squash", line 34). These linguistic expressions disambiguate the gestural form and point out that the conceptualization is in terms of physical manipulation, as if the action was performed

directly on the physical brain. On the other hand, the goal of the distortion correction is to manipulate the image, not its real-world referent.

It can be hypothesized that this coordination of different structures is given by the precise location of the gesture. The gesture is located in the space adjacent to the computer screen, i.e., the gesturing hand touches the screen, while, at the same time, it operates in the real world shared by the actors. This particular location of the gesture allows it to coordinate structures present in the digital image as well as in the three dimensional physical space. It enables us and the actors to read and identify the externalization of the integrated mental construct.

Notice, however, how relatively poor the individual representations are if taken in isolation. The image on the computer screen is completely static. The actors are dealing with a brain slice image that is simply displayed in the position that it assumes *before* and *after* computational manipulation. The linguistic utterances describe motion, but nothing is really moving on the computer screen (this is again an instance of fictive motion). The only element moving is the gesturing hand that touches the screen. The hand is shaped as if holding a three-dimensional object, but this object is absent from the action. Nevertheless, the conceptual structure achieved by the action is extremely rich and powerful. This richness and power is not achieved through any of the representations taken in isolation, but by their coordination. By selecting properties from each of the structures that partake in the action, and by their coordination, a complex structure - where actors at the same time communicate about and physically manipulate in front of them something that is two- and three-dimensional - is achieved.

It is obvious that this complex process directly depends on the form in which the fMRI data are presented, that is, visual images. Without the possibility of having visual images of the brain, the computation would have to be performed in a much more abstract and less efficient way (Arnheim, 1972; Schneiderman, 1983; Hutchins, Hollan, Norman, 1985). Instead, thanks to the visual form, the representations become embedded in real world situations of actions. Rather than dealing with some abstract representational form (e.g., numbers that the computer program manipulates), the object is dealt in a form that matches the way one thinks about the problem to be solved (Hutchins, Hollan, Norman, ibid.). Even though one knows that the process is computed through abstract mathematical means, one tends to think of the problem in terms of concrete, physical actions, e.g. rotation and squashing.

However, the analysis of coordination of various components of conceptual structure suggests that such cognitive process has an important element of fictivity in it. By coordinating the digital images with other representations (linguistic and gestural), as well as with the

---

[2] The example of such structures are watch, method of loci, or people standing in line. In the "standing in line conceptual network", an imagined trajector path is mapped onto a spatial array. This produces a sequentially ordered array in the blended space which gives meaning to the activity of standing in line. In this kind of network, conceptual relations are mapped onto relations among the material elements. One of the emergent properties of such construct is the stability of conceptual representation.

[3] Here I don't intend to imply an unidirectional path between internal mental states and its external counterpart in the gesture.

[4] Notice that the form of the image referent is deduced from the appearance of the digital image

space where the social action takes place, the actors are allowed to look at, and touch the impossible conceptual object. The fact that the researchers are dealing with a non-veridical conceptual structure can be inferred from their linguistic and gestural expressions. While such expressions refer to the brain and concrete physical actions (i.e., squashing and rotating), they also describe image correction (not its real-world referent), which is accomplished through abstract mathematical means. Hence, the concreteness of direct manipulation is achieved through the inclusion of the visual brain representations in the process that generates hybrid and non-veridical conceptual constructs.

Conclusion

By focusing on the details of scientific activity, my analysis suggests that visual representations play a fundamental role in conceptualization and knowledge acquisition in fMRI practice. Analysis of the collective action indicates that knowledge necessary for individuation of meaningful entities within a complex visual representation (i.e., retinotopic map) is constructed by recourse to materiality of the world, such as digital pictures of the brain activation, as well as bodily gestures and linguistic representations of both interlocutors. These different types of representations build cognitive constructs that, through conceptual compressions, transform abstract objects into manipulable, object-like entities.

The analyses above are consistent with the broader claim that conceptual structures are not confined to the individual mind. On the contrary, the social and material worlds are intrinsic elements of conceptualization and inference production (Goodwin, 1994; Hutchins, 1995). Not only in everyday, "simple" practices, but also in complex scientific activities, the external world enters into conceptualization to provide concreteness, that is, to allow scientists to think in terms of visual transformations, rather than in terms of abstract mathematical processes.

However, it is suggested that by looking at how visual images enter into coordination with other representational means, it is possible to see that the processes that generate concreteness are accomplished through inclusion of not only veridical, but also non-veridical and fictive elements. The observation of the activity of "retinotopy teaching" indicates that researchers employ non-veridical elements, such as fictive motion, in order to see and inscribe meaning (e.g., visual areas) to the brain image. Similarly, the analysis of the activity of distortion correction suggests that gesture, coordinated with the visual representation, language, and the use of space, manipulates an object that is at the same time physical and virtual.

References

Arnheim, R. (1972). *Visual Thinking*. Berkeley, CA: 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.

Coulson, S., & Oakley, T. (2000). "Blending Basics". *Cognitive Linguistics*, 11 –3/4.

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. (1998). "Conceptual Integration Networks", Cognitive Science, 22(2), 133-187.

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), 606-633.

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

Hutchins, E. Hollan, J., Norman, D. (1985). "Direct manipulation interfaces". *Human-computer interaction*, 1: 311-338.

Schneiderman, B. (1983). "Direct manipulation: A step beyond programming languages". IEEE Computer, 16(8), 57-69.

Talmy, L. (2000). *Toward a Cognitive Semantics*. Cambridge: MIT Press.