

# Computation and Agency in Scientific Cognition

Ronald N. Giere (giere@umn.edu)

Department of Philosophy & Center for Philosophy of Science  
University of Minnesota, Minneapolis, MN, USA 55455

## Abstract

I begin with a representative example of a contemporary scientific activity, observations using the Hubble Space Telescope, and ask what approaches within the cognitive sciences seem most fruitful as aids in developing an overall account of this sort of scientific activity. After presenting the Hubble Space Telescope System and a recent result, I consider applying a standard computational paradigm to this system. I find difficulties in identifying an appropriate cognitive agent and in making a suitable place for the instrumentation that constitutes such a large part of the whole system. I next consider applying the notion of distributed cognition as developed by Hutchins (1995), and then return to the question whether The Hubble System, understood as a distributed cognitive system, should be regarded as a computational system. I find a large computational component, but also an important part, the Hubble Telescope itself, that seems better characterized as a dynamic system than as a computational system. Moreover, the group of scientists interpreting the images produced by the system seem best thought of as a human/cultural system along the lines advocated by those developing a cognitive (Lakoff, 1987) or usage-based (Tomasello, 2003) approach to language acquisition and language use. I argue next that, while cognition may be theorized as distributed among both humans and instruments, there is no need to introduce into cognitive science a notion of distributed knowledge beyond simple collective knowledge. Even less is there any need to introduce notions of distributed mind or distributed consciousness. The result is that the agency involved in distributed cognitive systems remains simply human agency as ordinarily conceived. I conclude that distributed cognitive systems like The Hubble System are *hybrid* systems composed partly of dynamic physical systems, partly of computational systems, and partly of human cultural systems.

## Introduction

Science has many aspects, for example, individual, social, institutional and economic aspects. Focusing on the *activity* of doing science, it is undeniable that an important aspect of this activity is *cognitive*. In particular, the activity is supposed to produce new knowledge of the world. This must be counted as a cognitive activity, no matter what one's general characterization of "cognitive activities" might be.

From a disciplinary perspective, there are a number of ways one can approach the study of science as a cognitive activity. One way is to start from some area within the cognitive sciences and then to approach science as a subject matter to which one brings one's previously acquired methods and theories. Another way is to have been studying scientific activities from some other disciplinary

perspective, for example, the history, philosophy, or sociology science (collectively, science studies), and then to look to the cognitive sciences for resources with which to enrich such studies. This has always been my strategy.

A major difficulty for this latter strategy is that the cognitive sciences are themselves diverse, ranging from artificial intelligence, neuroscience, and cognitive psychology to cognitive anthropology. Moreover, both within and across these component disciplines there are major theoretical and methodological divides. If one is to produce a coherent account of scientific activities, one cannot avoid taking sides on controversies within the cognitive sciences themselves.

I begin with a representative example of a contemporary scientific activity and ask what approaches within the cognitive sciences seem most fruitful for developing a satisfactory account of the science. Historically, most attempts to exploit the resources of the cognitive sciences for studies of science have focused on individual scientific reasoning: inference, judgment, use of analogy, hypothesis formation, etc. (Gentner and Stevens, 1983; Nersessian, 1992). For the past several decades, interest in science studies has swung towards laboratory practices, including contemporary large-scale laboratories of the type found, for example, in high-energy physics, astronomy, genomics, and neuroimaging (Galison, 1997; Knorr-Cetina, 1999). Following this trend, I take my example from astronomy, in particular, the Hubble Space Telescope.

## The Hubble Space Telescope

The Hubble Space Telescope was launched on April 24, 1990 aboard the Space Shuttle Discovery. Figure 1 shows a schematic presentation of the Hubble Telescope.

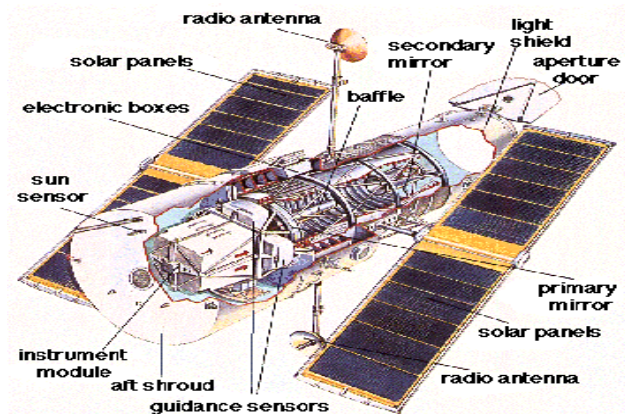


Figure 1. The Hubble Telescope.

Since a dramatic mission to correct an embarrassing flaw in its mirror at the end of 1993, Hubble has produced genuinely revolutionary observations. In January of 2003, for example, the Space Telescope Science Institute released a remarkable image produced by the Advanced Camera for Surveys (ACS) aboard the Hubble (which color image, I, unfortunately, cannot reproduce here).

Among several remarkable features of this particular image is that it involved gravitational lensing. During the exposure, the Hubble Telescope was pointed directly at a massive cluster of galaxies, known as Abell 1689, estimated to be 2.2 billion light-years away. In accordance with the General Theory of Relativity, this mass acts like a lens by warping space around it and thus effectively bending light passing by. Scientists who have studied the data claim that the image captures galaxies from which light was emitted roughly 13 billion years ago, when the universe was only one billion years old, which is to say, about one fourteenth its present age.

The process that produced this image is far too complex for me to describe here in any detail. In addition to the infamous mirror, it involves electronic detectors sensitive to light in the infrared part of the electro-magnetic spectrum. The output of the detectors is fed into an onboard computer and put into a form in which it can be transmitted to a Tracking and Data Relay Satellite from which it is retransmitted to the White Sands Complex near Las Cruces, New Mexico, from which it is again retransmitted by domestic satellite to the Data Operations Control Center at the Goddard Space Flight Center in Greenbelt, Maryland. From there it is routed to the Data Capture Facility and finally on to the Space Telescope Science Institute where it is studied by astronomers and other space scientists. The path from the data acquisition to production of the final image is shown in Figure 2. I will refer to this whole complex as the Hubble Space Telescope System (or just Hubble System), which includes the Hubble Space Telescope itself, the specific instrument launched in 1990, and depicted in Figure 1.

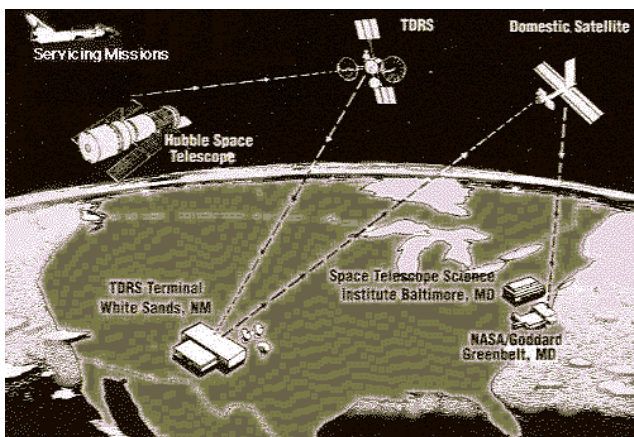


Figure 2. The path from data acquisition to final image.

To keep things manageable, let us take the cognitive output of this whole process to be just the claim that the image indicates the existence of galaxies 13 billion years ago.

### Understanding Observational Astronomy as a Cognitive Process

The general problem now is to understand the process leading to the conclusion about 13 billion old galaxies as a *cognitive* process. There is, of course, a standard cognitive science answer to our problem. It goes like this. There is a *cognitive agent*. This agent is an *individual* (human or artificial). This agent acquires a *symbolic representation* which is manipulated by *computation* according to definite *rules*. The general understanding is that these symbolic representations have the structure of a language and the rules the structure of a syntax. The endpoint of the computation is the conclusion about 13 billion old galaxies.

There are a number of difficulties that arise when one attempts to apply this paradigm to the Hubble System. One is locating the cognitive agent that acquires the representations and does the computations. The difficulty is not that there are no agents to be found. Rather, there seem to be too many agents. There is a whole team of people who control the movements of the telescope in space. Then there are whole teams of people at the Data Operations Control Center, the Data Capture Facility, and the Space Telescope Science Institute. And of course there are computers all over the place. One thing is clear. There is no one person that can be identified as *the* cognitive agent acquiring the representations and doing the computations.

One way of avoiding some of the problems just noted is to introduce the less standard notion of *collective cognition*, also sometimes called “socially shared cognition” (Resnick, Levine & Teasley, 1991). Understanding scientific cognition in terms of collective cognition eliminates the need to find a single cognitive agent. We can include as many participants as we like, whatever their role in the whole process. Moreover, the notion of collective cognition *can* be incorporated within the paradigm of cognition as computation. We need only assume that each individual is a computational system and that interactions among individuals are simply a matter of input and output restricted to individuals. The joint output can then be simply a combination of results computed within different individuals. The whole result is computed, only not within any one individual. Collective cognition, on this view, is fully computational. I will return to the issue whether collective cognition is best understood as fully computational. At the moment, there is another, more pressing issue.

The concept of collective cognition focuses exclusively on the human participants and neglects the undeniably large role played by instruments and other artifacts. The Hubble System is nothing without the Hubble Telescope itself. Nor would such a project be possible without high-speed computers to acquire and process gigabytes of data. This reflection takes us beyond simple collective cognition to a more general notion of *distributed cognition*.

## The *Palau* and The Hubble

As my source for the concept of distributed cognition, I take Ed Hutchins' (1995) ethnographic study of traditional "pilotage," that is, navigation near land as when coming into port. The cognitive task in this case is repeatedly determining the location of a traditional navy ship as it nears port. There are sailors on each side of the ship who telescopically record angular locations of landmarks relative to the ship's heading. These readings are then passed on via the ship's telephone to the pilothouse where they are combined by the navigator on a specially designed navigational chart on which bearings are drawn with a protractor-like device called a "hoey." The ship's position on the chart is determined by the intersection of two lines drawn using the bearings from the two sightings on opposite sides of the ship. Hutchins argues that parts of the cognitive process take place not in anyone's head, but in an instrument or on a chart. The cognitive process is distributed among humans and material artifacts.

It might seem farfetched to compare Hutchins' account of navigation aboard the *Palau* and the operations of The Hubble System. But everything depends on which aspects of the two systems are emphasized. In Hutchins' example, a number of people using various instruments produce a cognitive output that could not be produced by one person acting alone, nor by any number of people without the appropriate instruments. He argues that we should consider the whole complex, people acting in a particular social structure together with instruments, as a distributed cognitive system. It is this extended system that does the cognitive work. The Hubble Telescope System is similar in these respects. There is a social organization of people, together with instruments, that produces a cognitive output.

Whatever reasons there are for regarding determinations of the *Palau's* location as the product of a distributed cognitive system apply equally well for regarding the determination of the existence of galaxies in the universe 13 billion years ago as likewise the product of a distributed cognitive system. Of course there are also noteworthy differences. Determining the ship's position can be accomplished by a half dozen people using instruments originating in the 19th Century. The Hubble System requires hundreds, of people using sophisticated physical instruments and the latest computer technology.

Additionally, Hutchins' system is confined to the deck and pilothouse of a ship. The Hubble Telescope system extends at least from Earth orbit to the State of Maryland. If we count the lensing by the Abell 1689 galactic cluster as part of the system detecting the distant galaxies, the cognitive system is distributed 2.2 light years out into space. This raises in a dramatic way what Andy Clark (1997) calls the "identification problem." Once one allows that cognition may be taking place outside of someone's head (or outside the case of a computer), where are the boundaries? What counts as part of the cognitive system? Although this may seem like a serious problem if one insists on explicit criteria, there seems to me in practice little difficulty in deciding whether particular things are or are not playing a relevant causal role in the process of achieving the cognitive goal. It is abundantly clear that the Abell 1689 cluster played

an important causal role in producing the final image and the conclusions drawn from it. It must therefore be included in any account of this cognitive achievement. I will say no more about the individuation problem.

## But is the Hubble System Computational?

Is a distributed cognitive system a computational system? Hutchins (1995, 48) insists that "the activity of the larger system can be described in the way cognition has been traditionally described—that is, as computation realized through the creation, transformation, and propagation of representational states." He does so while acknowledging that the instruments used are best understood as analog to digital converters or vice versa. The spotting scope (alidade) converts analog position into angular coordinates. The hoey converts these coordinates back into an analog representation in the form of lines on a chart. Hutchins deliberately extends the notion of computation to cover these components of the system. He does so partly by invoking Marr's (1982) distinctions among the computational, representational or algorithmic, and implementational levels in a cognitive system. By abstracting to the computational level, one can temporarily forget about the fact that the implementation involves analog devices. I question whether this is a good move when thinking about large-scale experimental systems such as the Hubble Telescope System.

I would initially break the Hubble System down into three major components corresponding to the causal ordering of its interactions with the world. The first component is the front-end optical system. This consists of the Abell 1689 galactic cluster, the Hubble's mirror assembly, and the electronic sensors. The second component begins with the digital output of the sensors and ends with the production of the kind of images that appear on NASA's website. Each of these components is itself a distributed cognitive system including the hardware, software, and the many people who operate it. The third component consists of the scientific team that interprets the images and issues reports on their findings. This component is also a distributed cognitive system due to the interaction between the team members and the images plus data in other forms.

There is no questioning that the middle component includes a vast computational system in the strictest sense; a representation is created and transformed according to rule-governed programs. The output of the sensors is immediately digitized for transmission first to the Data Relay Satellite and eventually to the Goddard Space Flight Center in Maryland. Along the way, the representation is transformed in many ways thought to make it most informative to the astronomers who will eventually judge its scientific significance. The image we see at the end of this process is in no sense the data itself. It is better thought of as a *model* of the data. Even the output of the sensors is already a physically transformed form of the light that entered the telescope. Historians and sociologists of science are correct in claiming that the final image is, *to some extent*, a socially constructed object (Lynch and Woolgar, 1990).

It must be granted that straightforward computation is a large component in modern scientific cognition. Indeed, it is astounding the extent to which, since the 1950s, the computer has changed the way science is done. One could say that computers have changed the whole culture of science in fundamental ways. Nevertheless, it does not follow that the overall cognitive process of acquiring knowledge through the practice of experimental science is best characterized as computational. That remains to be seen.

### **Are Natural Physical Processes Computational?**

I say “natural” physical processes because an operating computer is a physical system. But it is a system that has been designed to instantiate a computational process. Few if any natural systems are like that. The natural system at issue in my example is the system including the Abel 1689 galactic cluster and the Hubble mirror and detectors. It is this system that gathers light from distant galaxies, directs it toward the sensors, and registers its intensity. Now, of course, at a very high level of abstraction, one can call this a computational system. If there are representations in the vicinity, they would presumably have to be a conjunction drawn from electromagnetic theory (for the light), optics (for the mirror), and solid state physics (for the detector, a silicon-based charge-coupled device). The algorithm would be provided by the associated mathematics. And the instantiation would be in an ordinary computer (pencil and paper having gone out of fashion). Is this helpful? I am not so sure.

Consider a much simpler example. Imagine dropping a solid ball from rest at a height  $h$  from the ground. This is a physical process in the same way that the operation of the front-end of the Hubble Telescope is a physical process. Is this process computational? Or, rather, are there good reasons for classifying this process as computational? I do not see that there are. There are no representations in this process. It is not representing anything. Nor are there any algorithms. There is no computing going on here. There is just the falling ball.

From mechanics we get the equation,  $y = h - \frac{1}{2} g t^2$ , for the height,  $y$ , of the falling ball as a function of time. Now, if we imagine the ball as somehow computing its distance from the ground, it would not be using this simple formula. This formula takes no account of air resistance or of the fact that the gravitational force of the earth is not constant, but varies with the distance from the center of the Earth. The ball, of course, must be sensitive to these and other factors influencing its fall. It is we, not falling balls, who compute the approximate distance from the ground using formulas like the one given above. In fact, no one ever has, and I doubt ever will, formulate a literally exact formula for even this simplest of mechanical systems. Better approximations, yes; literally exact values, no (see Giere 1988, Ch. 3; 1999, Ch. 5). The ball itself, of course, follows the exact values, whatever they might be.

I must admit, however, that these considerations are not completely decisive. Influential people in physics and computer science (notoriously, Wolfram 2002) argue that

the whole universe is one gigantic computer. They claim that space and time are discrete, so that it makes sense to say that the universe computes its next state (or the probability thereof) from its previous state. Now I agree with those (Lakoff & Johnson, 1980; Lakoff, 1987; Johnson 1987) who argue that language is deeply metaphorical. Nevertheless, in the sciences, the appropriateness of particular metaphors may be contested. In the 16<sup>th</sup> and 17<sup>th</sup> centuries, natural philosophers were much impressed with the elaborate clockworks then being produced. The universe, they concluded, is a gigantic clockwork. Newton’s physics was taken as providing legitimization for the literal truth of this claim. We now think this was mistaken. The digital computer is without doubt the dominant technological innovation of the current epoch. It is not surprising that it is taken as more than a mere metaphor, not only for the mind, but also for the universe as a whole. However compelling it now seems, I think this metaphor will turn out to be no better than the clockwork metaphor. I suspect that its attractiveness is at least partly rooted in a mistaken desire for a single, overarching explanation for everything. It is also a prime example of what Gerd Gigerenzer (2000) calls the move “from tools to theories,” as, indeed, is the cognition as computation paradigm itself.

There is a small minority in the cognitive sciences arguing that even humans should not be thought of as computational at all, but understood as dynamic systems (Thelen & Smith, 1994). At their most radical, they would do away with representations altogether in favor of dynamic systems. I would not go so far. I do, however, conclude that some parts of distributed cognitive systems are best thought of as dynamic, rather than computational, systems. Some parts of systems for experimentation in contemporary sciences seem to me good examples of such systems.

### **What About the Scientists?**

As noted earlier, it is possible to conceive of the scientists in my third component of the Hubble system as a system involving nothing more than distributed (collective) computation. Here the individual scientists would be regarded as computational systems using the produced images and other forms of data as inputs together with inputs from other agents. The collective output is the claim about 13 billion year old galaxies.

At this third stage we imagine a number of astronomers, astrophysicists, etc., examining the images and related forms of data, presumably comparing them with other images and data obtained at other times in other circumstances. Thinking of this combination of scientists and images as a distributed cognitive system, the cognitive process consists of the *interactions* among the scientists and their joint interactions with the images. The images are presumably manipulated (electronically, of course) in various ways, and the results of various manipulations discussed among members of the group. Whatever representations might be in the heads of the scientists, the most important representations are the *external* representations on the computer screens. Those are what are being evaluated for their implications regarding 13 billion year old galaxies. It seems to me initially as plausible to

regard all this activity as a dynamic system as it is to consider it to be a computational system. But neither of these interpretations seems to me right. Better to regard it as a human and cultural system. Here I can do no more than that point in the direction of support for this conclusion.

Broad support for a non-computational understanding of distributed cognitive systems dominated by humans comes from recent movements toward cognitive (Lakoff, 1987) and usage-based (Tomasello, 2003) linguistics. The leading idea is that learning a language is part of learning to be a member of a culture. This involves learning to interact with others as conscious intentional agents (Searle, 2002) and using language to get others to attend to particular objects in the environment and perform desired actions. It involves perception and motor skills as well as considerable knowledge about both the physical and human surroundings.

### Knowledge, Mind and Consciousness

If we say that new knowledge produced by the Hubble System is the product of a distributed cognitive system including both humans and artifacts, do we have to say that the resulting knowledge is likewise distributed among both humans and artifacts? Does distributed cognition require a distributed knower? And if a distributed knower, does it also require a distributed conscious mind, as some (Clark, 1997) have suggested?

Now I do not think these are factual questions. It is not as if anyone thinks they can design an experiment to detect a distributed mind. Rather, I think the issue is one of strategy in developing cognitive science as a science. For example, would it be fruitful for the development of cognitive science to introduce a notion of distributed minds to go along with the notion of distributed cognition? We are not forced by ordinary ways of thinking, or by philosophical tradition, or by the current state of cognitive science, to make such moves. We are free to make *cognition* a technical scientific concept different from everyday or traditional scientific notions. We can choose the theoretical principles of our science.

As for knowledge, there is a straightforward, though still theoretically interesting, way of restricting knowledge to the individual human components of a distributed cognitive system. The fact is that, in a complex system like the Hubble Telescope, the many people involved know different things. Thus, even though there may be more or less overlap in knowledge among members of any subgroup, the sum of everything known by someone is going to be much greater than what any single individual might know. In this way, knowledge is distributed among the people involved with operating the whole system. In the final stages of an experiment, only a few theoreticians may legitimately claim to know the result, for example, about galaxies 13 billion years ago. This relatively personal knowledge then becomes general scientific knowledge with the review, publication, and general consideration of the conclusion by the broader scientific community.

Once it is decided to restrict attribution of knowledge to individuals, whether singly or collectively, there is little motivation even to consider the possibility of distributed

minds, let alone distributed consciousness. So we end up with the cognitive process of acquiring new knowledge distributed among people and artifacts in a distributed cognitive system while knowledge, mind, and consciousness are restricted to individual people. In practice, this amounts to making cognition a specialized theoretical construct in cognitive science while keeping the concepts of knowledge, mind, and consciousness largely in their folk psychological states. The result is that the agency involved in distributed cognitive systems remains simply human agency as ordinarily conceived.

### Distributed Cognitive Systems as Hybrid Systems

The concept of a distributed cognitive system provides a unified way of understanding the large-scale experimental systems typical of contemporary science. At the same time, looking deeper into such systems reveals, I think, a *hybrid* system composed of further hybrid systems. By a hybrid system, I mean one that is partly a dynamic physical system, partly a computational system, and partly a human cultural system. In fact, each of the three components in my initial breakdown of the Hubble System is itself a hybrid system. The telescope requires humans and computers to position and operate it. The data stream from the satellite to the Goddard Space Flight Center is operated by many people. And the theoreticians at Goddard interact with their computer-generated images. This picture may lack the theoretical simplicity provided by the computational paradigm, but it is, I think, a better picture of the whole enterprise.

### Acknowledgements

The author wishes to acknowledge the support and hospitality of The Netherlands Institute for Advanced Study in the Humanities and Social Sciences (NIAS) and of the Fellows during the 2002-03 academic year.

### References

- Clark, A. (1997). *Being There: Putting Brain, Body, and World Together Again*. Cambridge, MA: MIT Press.
- Galison, P. L. (1997). *Image and Logic: A Material Culture of Microphysics*. Chicago, IL: University of Chicago Press.
- Gentner, D., & Stevens, A. L. (Eds.) (1983). *Mental Models*. Hillsdale, NJ: Erlbaum.
- Gigerenzer, G. (2000). *Adaptive Thinking*. New York: Oxford University Press.
- Giere, R. N. (1988). *Explaining Science: A Cognitive Approach*. Chicago, IL: University of Chicago Press.
- Giere, R. N. (1999). *Science Without Laws*. Chicago: University of Chicago Press.
- Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.
- Knorr-Cetina, K. (1999). *Epistemic Cultures: How the Sciences Make Knowledge*. Cambridge, MA: Harvard University Press.
- Johnson, M. (1987). *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago, IL: The University of Chicago Press.

- Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. Chicago, IL: University of Chicago Press.
- Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. Chicago, IL: University of Chicago Press.
- Lynch, M. & Woolgar, S. (Eds.) (1990). *Representation in Scientific Practice*. Cambridge, MA: MIT Press.
- Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco, CA: Freeman & Company.
- Nersessian, N. J. (1992). How Do Scientists Think? In R. N. Giere (Ed.), *Cognitive Models of Science* (Minnesota Studies in the Philosophy of Science, Vol. XV). Minneapolis, MN: University of Minnesota Press.
- Resnick, L. B., Levine, J. M., & Teasley, S. D. (Eds.) (1991). *Perspectives on Socially Shared Cognition*. Washington, DC: American Psychological Association.
- Searle, J. (2002). *Consciousness and Language*. Cambridge: Cambridge University Press.
- Thelen, E., & Smith, L. (1994). *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge, MA: MIT Press.
- Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.
- Wolfram, S. (2002). *A New Kind of Science*. Champaign, IL: Wolfram Media, Inc.