HOW WE BECAME POSTHUMAN

Virtual Bodies in Cybernetics, Literature, and Informatics

N. KATHERINE HAYLES

The University of Chicago Press
Chicago & London

N. KATHERINE HAYLES is professor of English at the University of California, Los Angeles. She holds degrees in both chemistry and English. She is the author of *The Cosmic Web: Scientific Field Models and Literary Strategies in the Twentieth Century* (1984) and *Chaos Bound: Orderly Disorder in Contemporary Literature and Science* (1990) and is the editor of *Chaos and Order: Complex Dynamics in Literature and Science* (1991), the last published by the University of Chicago Press.

The University of Chicago Press, Chicago 60637
The University of Chicago Press, Ltd., London

Printed in the United States of America
08 07 06 05 04 03 02 01 00 3 4 5

ISBN (cloth): 0-226-32145-2
ISBN (paper): 0-226-32146-0

Library of Congress Cataloging-in-Publication Data

Hayles, N. Katherine.
How we became posthuman : virtual bodies in cybernetics, literature, and informatics / N. Katherine Hayles.
p. cm.
Includes bibliographical references and index.
ISBN: 0-226-32145-2 (cloth : alk. paper). — ISBN: 0-226-32146-0 (pbk. : alk. paper)
1. Artificial intelligence. 2. Cybernetics. 3. Computer science. 4. Virtual reality. 5. Virtual reality in literature. I. Title.
Q335.H394 1999
003′.5—dc21 98-36459
CIP

♾ The paper used in this publication meets the minimum requirements of the American National Standard for the Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1992.

Contents

CONTESTING FOR THE BODY OF INFORMATION: THE MACY CONFERENCES ON CYBERNETICS

When and where did information get constructed as a disembodied medium? How were researchers convinced that humans and machines are brothers under the skin? Although the Macy Conferences on Cybernetics were not the only forum grappling with these questions, they were particularly important because they acted as a crossroads for the traffic in cybernetic models and artifacts. This chapter charts the arguments that made information seem more important than materiality within this research community. Broadly speaking, the arguments were deployed along three fronts. The first was concerned with the construction of information as a theoretical entity; the second, with the construction of (human) neural structures so that they were seen as flows of information; the third, with the construction of artifacts that translated information flows into observable operations, thereby making the flows "real."

Yet at each of these fronts, there was also significant resistance to the reification of information. Alternate models were proposed; important qualifications were voiced; objections were raised to the disparity between simple artifacts and the complex problems they addressed. Reification was triumphant not because it had no opposition but because scientifically and culturally situated debates made it seem a better choice than the alternatives. Recovering the complexities of these debates helps to demystify the assumption that information is more essential than matter or energy. Followed back to moments before it became a black box, this conclusion seems less like an inevitability and more like the result of negotiations specific to the circumstances of the U.S. techno-scientific culture during and immediately following World War II.

The Macy Conferences were unusual in that participants did not present finished papers. Rather, speakers were invited to sketch out a few main ideas to initiate discussion. The discussions, rather than the presentations, were the center of interest. Designed to be intellectual free-for-alls, the conferences were radically interdisciplinary. The transcripts show that researchers from a wide variety of fields—neurophysiology, electrical engineering, philosophy, semantics, literature, and psychology, among others—struggled to understand one another and make connections between others' ideas and their own areas of expertise. In the process, a concept that may have begun as a model of a particular physical system came to have broader significance, acting simultaneously as mechanism and metaphor.

The dynamics of the conferences facilitated this mixing. Researchers might not have been able to identify in their own work the mechanism discussed by a fellow participant, but they could understand it metaphorically and then associate the metaphor with something applicable to their own field. The process appears repeatedly throughout the transcripts. When Claude Shannon used the word "information," for example, he employed it as a technical term having to do with message probabilities. When Gregory Bateson appropriated the same word to talk about initiation rituals, he interpreted it metaphorically as a "difference that makes a difference" and associated it with feedback loops between contesting social groups. As mechanism and metaphor were compounded, concepts that began with narrow definitions spread out into networks of broader significance. Earlier I called these networks "constellations," suggesting that during the Macy period, the emphasis was on homeostasis. This chapter explores the elements that came together to form the homeostasis constellation; it also demonstrates the chain of associations that bound reflexivity together with subjectivity during the Macy period, which for many of the physical scientists was enough to relegate reflexivity to the category of "nonscience" rather than "science." Tracing the development of reflexive epistemologies after the Macy period ended, the chapter concludes by showing how reflexivity was modified so that it could count as producing scientific knowledge during the second wave of cybernetics.

The Meaning(lessness) of Information

The triumph of information over materiality was a major theme at the first Macy Conference. John von Neumann and Norbert Wiener led the way by making clear that the important entity in the man-machine equation was information, not energy. Although energy considerations are not entirely absent (von Neumann discussed at length the problems involved in dissipating the heat generated from vacuum tubes), the thermodynamics of heat was incidental. Central was how much information could flow through the system and how quickly it could move. Wiener, emphasizing the movement from energy to information, made the point explicitly: "The fundamental idea is the message . . . and the fundamental element of the message is the decision."[1] Decisions are important not because they produce material goods but because they produce information. Control information, and power follows.

But what counts as information? We saw in chapter 1 that Claude Shannon defined information as a probability function with no dimensions, no materiality, and no necessary connection with meaning. Although a full exposition of information theory is beyond the scope of this book, the following explanation, adapted from an account by Wiener, will give an idea of the underlying reasoning.[2] Like Shannon, Wiener thought of information as representing a choice. More specifically, it represents a choice of one message from among a range of possible messages. Suppose there are thirty-two horses in a race, and we want to bet on Number 3. The bookie suspects the police have tapped his telephone, so he has arranged for his clients to use a code. He studied communication theory (perhaps he was in one of the summer-school classes on communication theory that Wiener taught at UCLA), and he knows that any message can be communicated through a binary code. When we call up, his voice program asks if the number falls in the range of 1 to 16. If it does, we punch the number "1"; if not, the number "0." We use this same code when the voice program asks if the number falls in the range of 1 to 8, then the range of 1 to 4, and next the range of 1 to 2. Now the program knows that the number must be either 3 or 4, so it says, "If 3, press 1; if 4, press 0," and a final tap communicates the number. Using these binary divisions, we need five responses to communicate our choice.

How does this simple decision process translate into information? First let us generalize our result. Probability theory states that the number of binary choices C necessary to uniquely identify an element from a set with n elements can be calculated as follows:

$$C = \log_2 n$$

In our case,

$$C = \log_2 32 = 5,$$

the five choices we made to convey our desired selection. (Hereafter, to simplify the notation, consider all logarithms taken to base 2). Working from this formula, Wiener defined information I as the log of the number n of elements in the message set.

$$I = \log n$$

This formula gives I when the elements are equally likely. Usually this is not the case; in English, for example, the letter "e" is far more likely to occur than "z." For the more general situation, when the elements $s_1, s_2, s_3, \ldots s_n$ are not equally likely, and $p(s)$ is the probability that the element s will be chosen,

$$I(s_i) = \log 1/p(s_i) = -\log p(s_i).$$

This is the general formula for information communicated by a specific event, in our case the call to the bookie. Because electrical engineers must design circuits to handle a variety of messages, they are less interested in specific events than they are in the average amount of information from a source, for example, the average of all the different messages that a client might communicate about the horse race. This more complex case is represented by the following formula:

$$I = -\Sigma\, p(s_i)\,[\log p(s_i)],$$

where $p(s_i)$ is the probability that the message element s_i will be selected from a message set with n elements (Σ indicates the sum of terms as i varies from 1 to n).[3]

We are now in a position to understand the deeper implications of information as it was theorized by Wiener and Shannon. Note that the theory is formulated entirely without reference to what information means. Only the probabilities of message elements enter into the equations. Why divorce information from meaning? Shannon and Wiener wanted information to have a stable value as it moved from one context to another. If it was tied to meaning, it would potentially have to change values every time it was embedded in a new context, because context affects meaning. Suppose, for example, you are in a windowless office and call to ask about the weather. "It's raining," I say. On the other hand, if we are both standing on a street corner, being drenched by a downpour, this same response would have a very different meaning. In the first case, I am telling you something you don't know; in the second, I am being ironic (or perhaps moronic). An information concept that ties information to meaning would have to yield two different values for the two circumstances, even though the message ("It's raining") is the same.

To cut through this Gordian knot, Shannon and Wiener defined information so that it would be calculated as the same value regardless of the contexts in which it was embedded, which is to say, they divorced it from meaning. *In context,* this was an appropriate and sensible decision. *Taken out of context,* the definition allowed information to be conceptualized as if it were an entity that can flow unchanged between different material substrates, as when Moravec envisions the information contained in a brain being downloaded into a computer. Ironically, this reification of information is enacted through the same kind of decontextualizing moves that the theory uses to define information as such. The theory decontextualizes information; Moravec decontextualizes the theory. Thus, a simplification necessitated by engineering considerations becomes an ideology in which a reified concept of information is treated as if it were fully commensurate with the complexities of human thought.[4]

Shannon himself was meticulously careful about how he applied information theory, repeatedly stressing that information theory concerned only the efficient transmission of messages through communication channels, not what those messages mean. Although others were quick to impute larger linguistic and social implications to the theory, he resisted these attempts. Responding to a presentation by Alex Bavelas on group communication at the eighth Macy Conference, he cautioned that he did not see "too close a connection between the notion of information as we use it in communication engineering and what you are doing here . . . the problem here is not so much finding the best encoding of symbols . . . but, rather, the determination of the semantic question of what to send and to whom to send it."[5] For Shannon, defining information as a probability function was a strategic choice that enabled him to bracket semantics. He did not want to get involved in having to consider the receiver's mindset as part of the communication system. He felt so strongly on this point that he suggested Bavelas distinguish between information in a channel and information in a human mind by characterizing the latter through "subjective probabilities," although how these were to be defined and calculated was by no means clear.

Not everyone agreed that it was a good idea to decontextualize information. At the same time that Shannon and Wiener were forging what information would mean in a U.S. context, Donald MacKay, a British researcher, was trying to formulate an information theory that would take meaning into account. At the seventh conference, he presented his ideas to the Macy group. The difference between his view and Shannon's can be seen in the way he bridled at Shannon's suggestion about "subjective probabilities." In the rhetoric of the Macy Conferences, "objective" was associated with being scientific, whereas "subjective" was a code word implying that one had fallen into a morass of unquantifiable feelings that might be magnificent but were certainly not science. MacKay's first move was to rescue information that affected the receiver's mindset from the "subjective" label. He proposed that both Shannon and Bavelas were concerned with what he called "selective information," that is, information calculated by considering the selection of message elements from a set. But selective information alone is not enough; also required is another kind of information that he called "structural." Structural information indicates how selective information is to be understood; it is a message about how to interpret a message—that is, it is a metacommunication.

To illustrate, say I launch into a joke and it falls flat. In that case, I may resort to telling my interlocutor, "That's a joke." The information content of this message, considered as selective information (measured in "metrons"), is calculated with probability functions similar to those used in the Shannon-Wiener theory. In addition, my metacomment also carries structural information (measured in "logons"), for it indicates that the preceding message has one kind of structure rather than another (a joke instead of a serious statement). In another image MacKay liked to use, he envisioned selective information as choosing among folders in a file drawer, whereas structural information increased the number of drawers (jokes in one drawer, academic treatises in another).

Since structural information indicates how a message should be interpreted, semantics necessarily enters the picture. In sharp contrast to message probabilities, which have no connection with meaning, structural information was to be calculated through changes brought about in the receiver's mind. "It's raining," heard by someone in a windowless office, would yield a value for the structural information different from the value that it would yield when heard by someone looking out a window at rain. To emphasize the correlation between structural information and changes in the receiver's mind, MacKay offered an analogy: "It is as if we had discovered how to talk quantitatively about size through discovering its effects on the measuring apparatus."[6] The analogy implies that representations created by the mind have a double valence. Seen from one perspective, they contain information about the world ("It's raining"). From another perspective, they are interactive phenomena that point back to the observer, for this information is quantified by measuring changes in the "measuring instrument," that is, in the mind itself. And how does one measure these changes? An observer looks at the mind of the person who received the message, which is to say that changes are made in the observer's mind, which in turn can also be observed and measured by someone else. The progression tends toward the infinite regress characteristic of reflexivity. Arguing for a strong correlation between the *nature* of a representation and its *effect*, MacKay's model recognized the mutual constitution of form and content, message and receiver. His model was fundamentally different from the Shannon-Wiener theory because it triangulated between reflexivity, information, and meaning. In the context of the Macy Conferences, his conclusion qualified as radical: subjectivity, far from being a morass to be avoided, is precisely what enables information and meaning to be connected.

The problem was how to quantify the model. To achieve quantification, a mathematical model was needed for the changes that a message triggered in the receiver's mind. The staggering problems this presented no doubt explain why MacKay's version of information theory was not widely accepted among the electrical engineers who would be writing, reading, and teaching the textbooks on information theory in the coming decades. Although MacKay's work continued to be foundational for the British school of information theory, in the United States the Shannon-Wiener definition of information, not MacKay's, became the industry standard.

Not everyone in the United States capitulated. As late as 1968, Nicolas S. Tzannes, an information theorist working for the U.S. government, sent Warren McCulloch a memorandum about his attempt to revise MacKay's theory so that it would be more workable.[7] He wanted to define information so that its meaning varied with context, and he looked to Kotelly's context algebra for a way to handle these changes quantitatively. In the process, he made an important observation. He pointed out that whereas Shannon and Wiener define information in terms of what it *is*, MacKay defines it in terms of what it *does*.[8] The formulation emphasizes the reification that information undergoes in the Shannon-Wiener theory. Stripped of context, it becomes a mathematical quantity weightless as sunshine, moving in a rarefied realm of pure probability, not tied down to bodies or material instantiations. The price it pays for this universality is its divorce from representation. When information is made representational, as in MacKay's model, it is conceptualized as an action rather than a thing. Verblike, it becomes a process that someone enacts, and thus it necessarily implies context and embodiment. The price it pays for embodiment is difficulty of quantification and loss of universality.

In the choice between what information is and what it does, we can see the rival constellations of homeostasis and reflexivity beginning to take shape. Making information a thing allies it with homeostasis, for so defined, it can be transported into any medium and maintain a stable quantitative value, reinforcing the stability that homeostasis implies. Making information an action links it with reflexivity, for then its effect on the receiver must be taken into account, and measuring this effect sets up the potential for a reflexive spiral through an infinite regress of observers. Homeostasis won in the first wave largely because it was more manageable quantitatively. Reflexivity lost because specifying and delimiting context quickly ballooned into an unmanageable project. At every point, these outcomes are tied to the historical contingencies of the situation—the definitions offered, the models proposed, the techniques available, the allies and resources mobilized by contending participants for their views. Conceptualizing information as a disembodied entity was not an arbitrary decision, but neither was it inevitable.

The tension between reified models and embodied complexities figures importantly in the next episode of our story. If humans are information-processing machines, then they must have biological equipment enabling them to process binary code. The model constructing the human in these terms was the McCulloch-Pitts neuron. The McCulloch-Pitts neuron was the primary model through which cybernetics was seen as having "a setting in the flesh," as Warren McCulloch put it. The problem was how to move from this stripped-down neural model to such complex issues as universals in thought, gestalts in perception, and representations of what a system cannot represent. Here the slippage between mechanism and model becomes important, for even among researchers dedicated to a hard-science approach, such as McCulloch, the tendency was to use the model metaphorically to forge connections between relatively simple neural circuits and the complexities of embodied experience. In the process, the disembodied logical form of the circuit was rhetorically transformed from being an *effect* of the model to a *cause* of the model's efficacy. This move, familiar to us as the Platonic backhand, made embodied reality into a blurred and messy instantiation of the clean abstractions of logical forms. Unlike others who make this move, however, McCulloch never relinquished his commitment to embodiment. The tension between logical form and embodiment in his work displays how the construction of a weightless information was complicated when cybernetics moved into the intimate context of the body's own neural functioning.

Neural Nets as Logical Operators

Warren McCulloch figured large in the Macy Conferences. He chaired the meetings and, according to all accounts, was a strong leader who exercised considerable control over who was allowed to speak and who was not. He had studied philosophy under F. S. C. Northrop and was familiar with Rudolf Carnap's propositional logic. When he turned to neurophysiology, he was driven by two questions as much philosophical as scientific. "What is a number, that a man may know it, and a man, that he may know a number?"[9] He sought the answers in a model of a neuron that he envisioned as having two aspects—one physical, the other symbolic. The McCulloch-Pitts neuron, as it came to be called, was enormously influential. Although it has now been modified in significant ways, for a generation of researchers it provided the standard model of neural functioning. In its day, it represented a triumph of experimental work and theoretical reasoning. As Steve Heims points out, it was not easy to extrapolate from amorphous pink tissue on the laboratory table to the clean abstractions of the model.[10] Before complicating our story by looking at the interplay between logical form and complex embodiment, let us first consider the model on its own terms.

The McCulloch-Pitts neuron has inputs that can be either excitatory or inhibitory. A threshold determines how much excitation is needed for it to fire. A neuron fires only if the excitation of its inputs exceeds the inhibition by at least the amount of the threshold. Neurons are connected into nets. Each net has a set of inputs (signals coming in to neurons in the net), an output set (signals leading out from neurons in the net), and a set of internal states (determined by input, output, and signals from neurons that operate inside the net but are not connected to incoming or outgoing neurons). McCulloch's central insight was that neurons connected in this way are capable of signifying logical propositions. For example, if neurons A and B are connected to C and both are necessary for C to fire, this situation corresponds to the proposition, "If A and B are both true, then C is true." If either A or B can cause C to fire, the signified proposition is "If A or B is true, then C is true." If B is inhibitory and C will fire on input from A only if B does not fire, the signified proposition is "C is true only if A is true and B is not true." This much McCulloch had formulated by 1941 when he met Walter Pitts, a brilliant and eccentric seventeen-year-old who was to become his most important collaborator.[11] Pitts worked out the mathematics proving several important theorems about neural nets. In particular, he showed that a neural net can calculate any number (that is, any proposition) that can be calculated by a Turing machine.[12] The proof was important because it joined a model of human neural functioning with automata theory. Demonstrating that the operations of a McCulloch-Pitts neural net and a Turing machine formally converge confirmed McCulloch's insight "that brains do not secrete thought as the liver secretes bile but . . . they compute thought the way electronic computers calculate numbers."[13]

Although McCulloch knew as well as anyone that the McCulloch-Pitts neuron was a simplified schematic of an actual neuron's complexity, not to mention the brain's complexity, he pushed toward connecting the operations of a neural net directly with human thought. In his view, when a neuron receives an input related to a sensory stimulus, its firing is a direct consequence of something that happened in the external world. When he says a proposition calculated by a neural net is "true," he means that the event to which the firing refers really happened. How did McCulloch account for hallucinations and such phenomena as causalgia, an amputee's burning sensation that refers to a limb no longer present? He proposed that neural nets can set up reverberating loops that, once started, continue firing even though no new signals are incoming. To distinguish between firings signifying an external event and those caused by past history, he called the former "signals" and the latter "signs." A signal "always implies its occasion," but a sign is an "enduring affair which has lost its essential temporal reference."[14] The multiple meanings that McCulloch and his colleagues attached to reverberating loops indicate how quickly speculation leaped from the simplified model to highly complex phenomena. Lawrence Kubie linked reverberating loops with the repetitive and obsessive qualities of neuroses; numerous Macy participants suggested that the loops could account for gestalt perception; and McCulloch himself connected them not only with physical sensations but also with universals in philosophical thought.[15]

The gap between the relatively simple model and the complex phenomena it was supposed to explain is the subject of an exchange of letters between McCulloch and Hans-Lukas Teuber, a young psychologist who joined the Macy group on the fourth meeting and later became a coeditor of the published transcripts. Here, in correspondence with a junior colleague, McCulloch lays bare the assumptions that make embodied reality derivative from logical form. In a letter dated November 10, 1947, Teuber argues that similarity in outcome between different cybernetic systems does not necessarily imply similarity in structure or process. "Your robot may become capable of doing innumerable tricks the nervous system is able to do; it is still unlikely that the nervous system uses the same methods as the robot in arriving at what might look like identical results. Your models remain models—unless some platonic demon mediate between the investigators of organic structure and the diagram-making mathematicians." Only the psychologist, he claims, can give the neurophysiologist information on what "the most relevant aspects of the recipient structures [in sensory function] might be."[16] Cybernetic mechanisms do not signify unless they are connected with how perception actually takes place in human observers.

In his response on December 10, 1947, McCulloch explained his position. "I look to mathematics, including symbolic logic, for a statement of a theory in terms so general that the creations of God and man must exemplify the processes prescribed by that theory. Just because the theory is so general as to fit robot and man, it lacks the specificity required to indicate mechanism in man to be the same as mechanism in robot." In this argument, universality is achieved by bracketing or "black-boxing" the specific mechanisms. It emerges by erasing particularity and looking for general forms. Rhetorically, however, McCulloch presents the theory as though it *preexisted* specific mechanisms and then was later imperfectly instantiated in them. This backhanded swing invests the theory with a coercive power that cannot be ignored, for it expresses "a law so general" that "every circuit built by God or man must exemplify it in some form."[17]

In actuality, the theorem to which McCulloch refers is proved only in relation to the simplified model of a McCulloch-Pitts neural net. It therefore can have the coercive power he claims for it only if the assumptions made for the model also hold for embodied actuality, a congruence that can be exact only if the model is as complex and noisy as reality itself. Building such a model would, of course, defeat the purpose of model-making, as Lewis Carroll (and later Jorge Luis Borges) playfully points out when he imagines a king's mad cartographer who is satisfied only when he creates a map that covers the entire kingdom, reflecting its every detail in a scale of 1:1.[18] Teuber points to a gap when he ironically asks if some "platonic demon" is mediating between organic structure and abstract diagrams, a gap that has not been closed despite McCulloch's backhand volley.

In a feminist critique of the history of logic, Andrea Nye traces similar Platonic backhands that were made to develop a logic coercive in its lawlike power.[19] Nye points out that such moves are always made in specific political and historical contexts in which they have important social implications—implications that are masked by being presenting as preexisting laws of nature.[20] Like the logicians, McCulloch stripped away context to expose (or create) a universal form. But unlike the logicians, McCulloch in 1947 does not want to leave embodied reality behind. He is searching for an "empirical epistemology," a way of combining embodied actuality with the force of logical propositions. Teuber's objections hit a nerve (or neuron) because he insisted that the abstraction is not the actuality.

Dedicated to an empirical epistemology, McCulloch cannot rest content with interpreting logical form as a universal command that embodied flesh must obey. A suture is needed to bind the flesh more tightly to the model. The suture appears in his invocation of mechanisms that had previously been black-boxed in his appeal to universality. He recounts two instances when circuits he had sketched out for pattern-recognition in robots were identified by colleagues as accurate representations of the auditory and visual portions of the cortex—*in humans.* Now McCulloch—like a knight that, moved from the diagonal to attack the queen, exposes the queen to the bishop's attack as well—has caught Teuber in a two-pronged attack. In the first approach, humans and robots are judged alike because they obey the same universal law, whatever their mechanisms. In the second approach, humans and robots are judged alike because they use the same mechanisms. This double attack is also invoked, as we shall see in the next chapter, by Norbert Wiener and his collaborators when a young upstart philosopher took issue with their cybernetic manifesto. It tends to appear when cybernetic arguments are challenged because it allows a defense on two fronts simultaneously. If mechanisms are black-boxed so that only behavior counts, humans and robots look the same because they (can be made to) behave the same. If the black boxes are opened up (and viewed from carefully controlled perspectives), the mechanisms inside the boxes look the same, again demonstrating the equivalence.

How can the queen be saved? By recognizing that the abstractions here are multilayered. When McCulloch goes down a level, away from what information is toward what it does, he still ends up several layers away from embodied complexity. Consider his claim that pattern-recognition circuits in a robot mechanism and in a human cortex are the same. These circuits are diagrams that have been abstracted from two different kinds of embodiments, neural tissue for the human and vacuum tubes or silicon chips for the robot. Although there may be a level of abstraction at which similarities can be made to appear, there is also a level of specificity at which differences create a significant gap. It depends on how the perspective is constructed. Controlling the context, particularly the movement from instantiated specificity to abstraction, was crucial to constructing the pathways through which the McCulloch-Pitts neuron was made to stand simultaneously for a computer code and for human thought. Transforming the body into a flow of binary code pulsing through neurons was an essential step in seeing human being as an informational pattern. *In context,* this transformation can be seen as a necessary simplification that made an important contribution to neurophysiology. *Taken out of context,* it is extrapolated to the unwarranted conclusion that there is no essential difference between thought and code.

I admire McCulloch because he made the audacious leap from amorphous tissue to logical model; I admire him even more because he resisted the leap. Although he emphasized the ability of his neurons to formulate propositions, he never saw them as disembodied. He was aware that information moves only through signals and that signals have existence only if they are embodied. "By definition, a signal is a proposition embodied in a physical process," he asserted in a speech, entitled "How Nervous Structures Have Ideas," to the American Neurological Association in 1949.[21] In the context of his writing as a whole, a commitment to embodiment exists in dynamic tension with an equally strong proclivity to see embodiment as the instantiation of abstract propositions.

This tension can be seen in the manuscript version of "What's in the Brain That Ink May Character?" dated August 28, 1964. McCulloch recounts about a recent trip to Ravello: "I was told that an automaton or a nerve net, like me, was a mapping of a free monoid onto a semigroup with the possible addition of identity." The parenthetical "like me" points up the incongruity between a highly abstract mathematical model involving monoids and semigroups and the embodied creature who pens these lines. "This is the same sort of nonsense one finds in the writings of those who never understood [abstract form] as an embodiment," he continues. "It is like mistaking a Chomsky language for a real language. You will find no such categorical confusion in the original Pitts and McCulloch of 1943. There the temporal propositional expressions are events occurring in time and space in a physically real net. The postulated neurons, for all their oversimplifications, are still physical neurons as truly as the chemist's atoms are physical atoms."[22] Here, in the slippages between abstract propositions, models of neurons, and "physically real" nets, we can see McCulloch trying to keep three balls in the air at once. Although the neurons are only "postulated" and are admittedly "oversimplifications," McCulloch fiercely wants to insist they are still physical. If he does not entirely succeed in creating an "empirical epistemology," he nevertheless achieves no small feat in insisting that none of the balls can be dropped without sacrificing the complexities of embodied thought.

The McCulloch-Pitts neuron is a liminal object, part abstraction and part embodied actuality, but other models were more firmly in the material realm. Part of what made cybernetics convincing to Macy participants and others were the electromechanical devices that showed cybernetic principles in action. Cybernetics was powerful because it *worked.* If you don't believe, watch William Grey Walter's robot tortoise returning to its cage for an electrochemical nip when its batteries are running low, or see Wiener's

Moth turning to follow the light and his Bedbug scuttling under a chair to avoid it. These devices were simple mechanisms by contemporary standards. Nevertheless, they served an important function because they acted as material instantiations of the momentous conclusion that humans and robots are siblings under the skin. Particularly important for the Macy Conferences were Shannon's electronic rat, a goal-seeking machine that modeled a rat learning a maze, and Ross Ashby's homeostat, a device that sought to return to a steady state when disturbed. These artifacts functioned as exchangers that brought man and machine into equivalence; they shaped the kinds of stories that participants would tell about the meaning of this equivalence. In conjunction with the formal theories, they helped to construct the human as cyborg.

The Rat and the Homeostat: Looping between Concept and Artifact

There are moments of clarity when participants came close to explicitly articulating the presuppositions informing the deep structure of the discussion. At the seventh conference, John Stroud, of the U.S. Naval Electronic Laboratory in San Diego, pointed to the far-reaching implications of Shannon's construction of information through the binary distinction between signal and noise. "Mr. Shannon is perfectly justified in being as arbitrary as he wishes," Stroud observed. "We who listen to him must always keep in mind that he has done so. Nothing that comes out of rigorous argument will be uncontaminated by the particular set of decisions that were made by him at the beginning, and it is rather dangerous at times to generalize. If we at any time relax our awareness of the way in which we originally defined the signal, we thereby automatically call all of the remainder of the received message the 'not' signal or noise."[23] As Stroud realized, Shannon's distinction between signal and noise had a conservative bias that privileges stasis over change. Noise interferes with the message's exact replication, which is presumed to be the desired result. The structure of the theory implied that change was deviation and that deviation should be corrected. By contrast, MacKay's theory had as its generative distinction the difference in the state of the receiver's mind before and after the message arrived. In his model, information was not opposed to change; it was change.

Applied to goal-seeking behavior, the two theories pointed in different directions. Privileging signal over noise, Shannon's theory implied that the goal was a preexisting state toward which the mechanism would move by making a series of distinctions between correct and incorrect choices. The goal was stable, and the mechanism would achieve stability when it reached the goal. This construction easily led to the implication that the goal, formulated in general and abstract terms, was less a specific site than stability itself. Thus the construction of information as a signal/noise distinction and the privileging of homeostasis produced and were produced by each other. By contrast, MacKay's theory implied that the goal was not a fixed point but was a changing series of values that varied with context. In his model, setting a goal temporarily marked a state that itself would become enfolded into a reflexive spiral of change. In the same way that signal/noise and homeostasis went together, so did reflexivity and information as a signifying difference.

These correlations imply that before Shannon's electronic rat ever set marker in maze, it was constituted through assumptions that affected how it would be interpreted. Although Shannon called his device a maze-solving machine, the Macy group quickly dubbed it a rat.[24] The machine consisted of a five-by-five square grid, through which a sensing finger moved. An electric jack that could be plugged into any of the twenty-five squares marked the goal, and the machine's task was to move through the squares by orderly search procedures until it reached the jack. The machine could remember previous search patterns and either repeat them or not, depending on whether they had been successful. Although Heinz von Foerster, Margaret Mead, and Hans Teuber—in their introduction to the eighth conference volume—highlighted the electronic rat's significance, they also acknowledged its limitations. "We all know that we ought to study the organism, and not the computers, if we wish to understand the organism. Differences in levels of organization may be more than quantitative." They go on to argue, however, that "the computing robot provides us with analogs that are helpful as far as they seem to hold, and no less helpful whenever they break down. To find out in what ways a nervous system (or a social group) differs from our man-made analogs requires experiment. These experiments would not have been considered if the analog had not been proposed."[25]

There is another way to understand this linkage. By suggesting certain kinds of experiments, the analogs between intelligent machines and humans *construct the human in terms of the machine.* Even when the experiment fails, the basic terms of the comparison operate to constitute the signifying difference. If I say a chicken is not like a tractor, I have characterized the chicken in terms of the tractor, no less than when I assert that the two are alike. In the same way, whether they are understood as like or unlike, ranging human intelligence alongside an intelligent machine puts the two into a relay system that constitutes the human as a special kind of information machine and the information machine as a special kind of human.[26] Although some characteristics of the analogy may be explicitly denied, the basic linkages it embodies cannot be denied, for they are intrinsic to being able to think the model. Presuppositions embodied in the electronic rat include the idea that both humans and cybernetic machines are goal-seeking mechanisms that learn, through corrective feedback, to reach a stable state. Both are information processors that tend toward homeostasis when they are functioning correctly.

Given these assumptions, it was perhaps predictable that reflexivity should be constructed as neurosis in this model. Shannon, demonstrating how his electronic rat could get caught in a reflexive loop that would keep it circling endlessly around, remarked, "It has established a vicious circle, or a singing condition."[27] "Singing condition" is a phrase that Warren McCulloch and Warren Pitts had used, in an earlier presentation, to describe neuroses modeled through cybernetic neural nets. If machines are like humans in having neuroses, humans are like machines in having neuroses that can be modeled mechanically. Linking humans and machines in a common circuit, the analogy constructs both of them as steady state systems that become pathological when they fall into reflexivity. This kind of mutually constitutive interaction belies the implication, inscribed in the volume's introduction, that such analogs are neutral heuristic devices. More accurately, they are relay systems that transport assumptions from one arena to the next.[28]

The assumptions traveling across the relay system set up by homeostasis are perhaps most visible in the discussion of W. Ross Ashby's homeostat.[29] The homeostat was an electrical device constructed with transducers and variable resistors. When it received an input changing its state, it searched for the configuration of variables that would return it to its initial condition. Ashby explained that the homeostat was meant to model an organism which must keep essential variables within preset limits to survive. He emphasized that the cost of exceeding those limits is death. If homeostasis equals safety ("Your life would be safe," Ashby responded when demonstrating how the machine could return to homeostasis), departure from homeostasis threatens death (p. 79). One of his examples concerns an engineer sitting at the control panel of a ship. The engineer functions like a homeostat, striving to keep the dials within certain limits to prevent catastrophe. Human and machine are alike in needing stable interior environments. The human keeps the ship's interior stable, and this stability preserves the homeostasis of the human's interior, in turn allowing the human to continue to ensure the ship's homeostasis. Arguing that homeostasis is a requirement "uniform among the inanimate and the animate," Ashby privileged it as a universally desirable state (p. 73).

The postwar context for the Macy Conferences played an important role in formulating what counted as homeostasis. Given the cataclysm of the war, it seemed self-evident that homeostasis was meaningful only if it included the environment as part of the picture. Thus Ashby conceived of the homeostat as a device that included both the organism and the environment. "Our question is how the organism is going to struggle with its environment," he remarked, "and if that question is to be treated adequately, we must assume some specific environment" (pp. 73–74). This specificity was expressed through the homeostat's four units, which could be arranged in various configurations to simulate organism-plus-environment. For example, one unit could be designated "organism" and the remaining three the "environment"; in another arrangement, three of the units might be the "organism," with the remaining one the "environment." Formulated in general terms, the problem the homeostat addressed was this: given some function of the environment E, can the organism find an inverse function E^{-1} such that the product of the two will result in a steady state? When Ashby asked Macy participants whether such a solution could be found for highly nonlinear systems, Julian Bigelow correctly answered, "In general, no" (p. 75). Yet, as Walter Pitts observed, the fact that an organism continues to live means that a solution does exist. More precisely, the problem was whether a solution could be articulated within the mathematical conventions and technologies of representation available to express it. These limits in turn were constituted through the model's specificities that translated between the question in the abstract and the particular question posed by that experiment. Thus the emphasis shifted from finding a solution to stating the problem.

This dynamic appears repeatedly throughout the Macy discussions. Participants increasingly understood the ability to specify exactly what was wanted as the limiting factor for building machines that could perform human functions. Von Neumann stated the thesis at the first conference, and Walter Pitts restated it near the end of the meetings, at the ninth conference. "At the very beginning of these meetings," Pitts recalled, "the question was frequently under discussion of whether a machine could be built which would do a particular thing, and, of course, the answer, which everybody has realized by now, is that as long as you definitely specify what you want the machine to do, you can, in principle, build a machine to do it" (p. 107). After the conferences were over, McCulloch repeated this dynamic in *Embodiments of Mind.* Echoing across two decades, the assertion has important implications for language.

If what is exactly stated can be done by a machine, the residue of the uniquely human becomes coextensive with the linguistic qualities that interfere with precise specification—ambiguity, metaphoric play, multiple encoding, and allusive exchanges between one symbol system and another. The uniqueness of human behavior thus becomes assimilated to the ineffability of language, and the common ground that humans and machines share is identified with the univocality of an instrumental language that has banished ambiguity from its lexicon. Through such "chunking" processes, the constellations of homeostasis and reflexivity assimilated other elements into themselves. On the side of homeostasis was instrumental language, whereas ambiguity, allusion, and metaphor stood with reflexivity.

By today's standards, Ashby's homeostat was a simple machine, but it had encoded within it a complex network of assumptions. Paradoxically, the model's simplicity facilitated the overlay of assumptions onto the artifact, for its very lack of complicating detail meant that the model stood for much more than it physically enacted. During discussion, Ashby acknowledged that the homeostat was a simple model and asserted that he "would like to get on to the more difficult case of the clever animal that has a lot of nervous system and is, nevertheless, trying to get itself stable" (p. 97). The slippage between the simplicity of the model and the complexity of the phenomena did not go unremarked. J. Z. Young, from the Anatomy Department at University College, London, sharply responded: "Actually that is experimentally rather dangerous. You are all talking about the cortex and you have it very much in mind. Simpler systems have only a limited number of possibilities" (p. 100). Yet the "simpler systems" helped to reinforce several ideas: humans are mechanisms that respond to their environments by trying to maintain homeostasis; the function of scientific language is exact specification; the bottleneck for creating intelligent machines lies in formulating problems exactly; and an information concept that privileges exactness over meaning is therefore more suitable to model construction than one that does not. Ashby's homeostat, Shannon's information theory, and the electronic rat were collaborators in constructing an interconnected network of assumptions about language, teleology, and human behavior.[30]

These assumptions did not go uncontested. The concept that most clearly brought them into question was reflexivity. As we have seen, during the Macy Conferences reflexivity was a nebulous cluster that was not explicitly named as such. To give the flavor of the discussions that both invoked the possibility of reflexivity and failed to coalesce into coherent theory about it, we can consider the image of the man-in-the-middle. The image was given currency by World War II engineering technologies that aimed to improve human performance by splicing humans into feedback loops with machines. The image takes center stage in the sixth conference during John Stroud's analysis of an operator sandwiched between a radar-tracking device on one side and an antiaircraft gun on the other. The gun operator, Stroud observed, is "surrounded on both sides by very precisely known mechanisms and the question comes up, 'What kind of a machine have we put in the middle?' "[31] The image as Stroud used it constructs the man as an input/output device. Information comes in from the radar, travels through the man, and goes out through the gun. The man is significantly placed in the *middle* of the circuit, where his output and his input are already spliced into an existing loop. Were he at the end, it might be necessary to consider more complex factors, such as how he was interacting with an open-ended and unpredictable environment. The focus in Stroud's presentation was on how information is transformed as it moves through the man-in-the-middle. As with the electronic rat and the homeostat, the emphasis was on predictability and homeostatic stability.

Countering this view was Frank Fremont-Smith's insistence on the observer's role in constructing the image of the man-in-the-middle. "Probably man is never only between the two machines," he pointed out. "Certainly he is never only in between two machines when you are studying him because you are the other man who is making an input into the man. You are studying and changing his relation to the machines by virtue of the fact that you are studying him." Fremont-Smith's introduction of the observer was addressed by Stroud in a revealing image that sought to convert the observer into a man-in-the-middle. "The human being is the most marvelous set of instruments," Stroud observed, "but like all portable instrument sets the human observer is noisy and erratic in operation. However, if these are all the instruments you have, you have to work with them until something better comes along."[32] In Stroud's remark, the man is converted from an open-ended system into a portable instrument set. The instrument may not be physically connected to two mechanistic terminals, the image implied, but this lack of tight connection only makes the splice invisible. It does not negate the suture that constructs the human as an information-processing machine spliced into a closed circuit that ideally should be homeostatic in its operation, however noisy it is in practice.

Fremont-Smith responded: "You cannot possibly, Dr. Stroud, eliminate the human being. Therefore what I am saying and trying to emphasize is that, with all their limitations, it might be pertinent for those scientific investigators at the general level, who find to their horror that we have to work with human beings, to make as much use as possible of the insights available as to what human beings are like and how they operate."[33] As his switch to formal address indicates, Fremont-Smith was upset at the recuperation of his comment back into the ideology of objectivism. His comment cuts to the heart of the objection against reflexivity. Just as with MacKay's model of structural information, reflexivity opens the man-in-the-middle to psychological complexity, so that he can no longer be constructed as a black box functioning as an input/output device. The fear is that under these conditions, reliable quantification becomes elusive or impossible and science slips into subjectivity, which to many conferees meant that it was not real science at all. Confirming traditional ideas of how science should be done in a postwar atmosphere that was already clouded by the hysteria of McCarthyism, homeostasis implied a return to normalcy in more than one sense.

The thrust of Fremont-Smith's observations was, of course, to intimate that psychological complexity was unavoidable. The responses of other participants reveal that this implication was precisely what they were most concerned to deny. They especially disliked reflexive considerations that took the personal form of suggesting that their statements were not assertions about the world but were revelations of their own internal states. The primary spokesperson for this disconcerting possibility was Lawrence Kubie, a psychoanalyst from the Yale University Psychiatric Clinic. In correspondence, Kubie enraged other participants by interpreting their criticisms of his theories as evidence of their subconscious resistances rather than as matters for scientific debate. In his presentations he was more tactful, but the reflexive thrust of his arguments remained clear. His presentations occupy more space in the published transcripts than those of any other participant, composing about one-sixth of the total. Although he met with repeated skepticism among the physical scientists, he continued to defend his position. At the center of his explanation was the multiply encoded nature of language, which operated at once as an instrument that the speaker could use to communicate and as a reflexive mirror that revealed more than the speaker knew. Like MacKay's theory of information, Kubie's psychoanalytic approach built reflexivity into the model. Also like MacKay's theory, Kubie's argument met the greatest (conscious?) resistance in the demand for reliable quantification.

Kubie's ideas will serve as a springboard for looking at the role that reflexivity played in the Macy Conferences and in the lives of some participants after the conferences ended, particularly the lives of Margaret Mead and Gregory Bateson and their daughter, Mary Catherine Bateson. Contrasting the Macy Conferences with Catherine Bateson's account of a similar conference held in 1968 will illustrate why the full implications of reflexivity could scarcely have been admitted during the Macy period. Once the observer is made a part of the picture, cracks in the frame radiate outward until the perspectives that controlled context are fractured as irretrievably as a safety-glass windshield hit by a large rock. The Macy participants were right to feel wary about reflexivity. Its potential was every bit as explosive as they suspected.

Kubie's Last Stand

Lawrence Kubie had been trained as a neurophysiologist. He won McCulloch's admiration for his 1930 paper suggesting that neuroses were caused by reverberating loops similar to those McCulloch later modeled in neural nets.[34] In midcareer Kubie converted to psychoanalysis. By the time of the Macy Conferences, he was affiliated with the hard-line Freudianism of the New York Psychoanalytic Institute. In his presentation at the sixth conference, he laid out the fundamentals of his position. Neurotic processes are dominated by unconscious motivations. As goal-seeking behavior, these processes are ineffective because the unconscious pursues its goals in symbolic form. A man wants to feel secure, and money symbolizes this security for him. But when he acquires money, he still does not feel secure. He has acquired the symbol but lacks what the symbol represents. With the gap between desire and reality yawning as widely as ever, he may actually feel more rather than less anxious as he approaches his putative goal.

Although McCulloch thought of Kubie as an experimentalist, from the beginning of the conferences Kubie resisted the reductive approach that was characteristic of McCulloch's work. At the first conference, Kubie expressed uneasiness over reducing complex psychological phenomena to mechanistic models equating humans and automata. At the sixth conference he was still resisting. In "Neurotic Potential and Human Adaptation," he explained why he had not addressed feedback mechanisms: "I wanted to make clear the complexity and subtlety of the neurotic process as it is encountered clinically. Without this we are constantly in danger of oversimplifying the problem so as to scale it down for mathematical treatment."[35] Instead of mechanistic models, his formulations emphasized the reflexivity of psychological processes. At the seventh conference, in "The Relation of Symbolic Function in Language Formation and in Neurosis," he insisted on "the fact that the human organism has two symbolic functions and not one. One is language. The other is neurosis." Moreover, the two functions converge into the same utterance. Fremont-Smith drove the point home.

"What Dr. Kubie is really trying to say is that language is a double coding: both a statement about the outside and a statement about the inside. It is that doubleness which gives this conscious/unconscious quality to it."[36]

In this view, a statement intended as an observation of the external world is pierced by reflections of the speaker's interior state, including neurotic processes of which the speaker is not conscious. If a scientist denies this is the case, insisting that he or she speaks solely about external reality, these objections themselves can be taken as evidence of unconscious motivations. For experimentalists like McCulloch, concerned to give an objective account of mental processes, psychoanalysis was the devil's plaything because it collapsed the distance between speaker and language, turning what should be scientific debate into a tar baby that clung to them the more they tried to push it away.

The damage that this view of reflexive utterance could do to scientific objectivity was dramatically laid out by McCulloch in a 1953 address to the Chicago Literary Club. Entitled "The Past of a Delusion," the speech was a fiery denunciation of Freudian psychoanalysis.[37] If all scientific utterance is tinged with subjectivity, McCulloch felt, then scientific theory must inextricably be tied to the foibles and frailties of humans as subjective beings. To show the disastrous effects that this close coupling could have on science, McCulloch took as his case study Freudian psychoanalysis, a theory that in his view both promoted the idea of close coupling and itself insidiously instantiated it. McCulloch ripped into Freud, suggesting that Freud had turned to psychoanalysis because he had wanted to make more money than he would have as a Jewish medical doctor. McCulloch recounted Freud's sex life, intimating that Freud put sexuality at the heart of his theory because he was sexually frustrated himself. McCulloch denounced psychoanalysts as charlatans who, motivated by greed, kept treating their patients as long as those patients had money to pay. He sneered at the empirical evidence used by Freud and other psychoanalysts. In his ironic conclusion, McCulloch cautioned his audience not to try to argue with psychoanalysts. All they would get for their pains, he predicted, were psychoanalytic interpretations of their objections as evidence of their own unconscious hostilities.

Kubie learned of this speech from a colleague who had been in the audience.[38] Although McCulloch went out of his way to exempt Kubie from his general scorn for psychoanalysis (in a 1950 letter to Fremont-Smith, he had written, "Of all the psychoanalysts I know, [Kubie] has the clearest head for theory"),[39] the attack was too stinging not to draw a rejoinder. As pat as McCulloch would have wished, Kubie interpreted the speech as a sign of

McCulloch's own psychological distress. Speaking to a colleague, Kubie noted that McCulloch's "vitriole may be due to an accumulation of personal frustrations of his own displaced onto analysis."[40] Later, when he heard about McCulloch's erratic behavior during a presentation at Yale, he wrote to McCulloch's host, sending a copy of the letter to Fremont-Smith: "I am distressed by this news about Warren . . . in him the boundary between sickness and health has always been narrow" (p. 137). Kubie even tried to arrange for psychoanalysts in the Boston area to meet with McCulloch "on a social pretext if necessary," with a view to getting him the "help" that Kubie thought he needed (p. 138). As Steve Heims observes in his account of these incidents, McCulloch would have been enraged had he known about Kubie's attempts at intervention.

McCulloch's "The Past of a Delusion" is vivid evidence that Fremont-Smith's attempts at reconciliation between psychoanalysts and physical scientists did not succeed. Kubie was well aware of the experimentalists' attitudes. After repeated attempts to win them over, he delivered his final presentation at the ninth conference in what sounds like a state of controlled rage. He likened the supposed "troublemaker" psychiatrist to "a naturalist, reporting on the facts of human nature as observed by him." By contrast, he noted, the physical scientists ignore complex psychological phenomena in favor of the simplifications of an abstract model. "The experimentalist and mathematician then offer their explanation, whereupon, the naturalist presents additional observations which confront the experimentalist and the mathematician with an even more complex version of natural phenomena." As the cycle continues, "these new complexities are accepted with increasing reluctance and skepticism."[41] In these remarks Kubie presented his version of his presentations at the Macy Conferences. He merely reported on the facts, whereas the others offered inadequate mechanistic explanations for them. This characterization ignores, of course, the Freudian framework he used to interpret his colleagues' behavior, a framework at least as theory-laden in its observations as anything McCulloch proposed.[42]

I think of this presentation, loaded with controlled anger as if in point/counterpoint to McCulloch's extravagant display of anger in his speech of the following year, as Kubie's last stand. The resistance it describes and inscribes went in both directions, from the psychoanalyst to the experimentalist and from the experimentalist to psychoanalyst. For the experimentalists, psychoanalysis strengthened the chain of association that bound reflexivity together with subjectivity, for it added to the already daunting problems of quantification the unfalsifiable notion of the unconscious. It is no wonder that reflexivity came to seem, for many of the participants, a dead end for legitimate scientific inquiry.

Even as one version of reflexivity fizzled out, other versions were being constructed in terms that made them more productive, in part because these versions avoided associating reflexivity with the unconscious. Temple Burling, reading the published transcripts in 1954, wrote to McCulloch: "I was surprised at the jamb that the group got into at this late date over the question of 'the unconscious.' It seems to me that is putting the cart before the horse. It isn't unconscious neuro activity that is puzzling but conscious. Consciousness is the great mystery."[43] Burling's comments point to another way into reflexivity, a way taken by a handful of participants, including Heinz von Foerster, Margaret Mead, and Gregory Bateson. Though they were not necessarily opposed to psychoanalytic interpretation, it was not the focus of their attention. The scale on which they wanted to play their tunes did not run up and down the conscious/unconscious keyboard. Rather, they wanted to create models that would take into account the observer's role in constructing the system. The important dichotomy for them was observer/system, and the important problems were how to locate the observer inside the system and the system inside the observer.

Circling the Observer

In 1969, near the end of his career, Fremont-Smith wrote (or rather, had his secretary write) to participants of the various Macy Conferences that he had organized over three decades, asking for their evaluation of the interdisciplinary programs and the discussion formats. The inquiry was clearly a career-closing move; he was looking for affirmation of what he considered his lifework. Some of the replies were disarmingly frank. Jimmie Savage wrote about how it felt to be a young man allowed to "hobnob with such a diverse group of illustrious and brilliant people." He recalled that he had frequently found himself thinking that the emperor had no clothes but wondering if he could trust his own feelings. He confessed, "Cybernetics itself seemed to me to be mostly baloney."[44] R. W. Gerard expressed similar dissatisfactions, recalling being "intensely frustrated by the perpetual tangents to tangents that developed during a meeting and the rare satisfaction of intellectual closure and completion of any line of thought or argument." He added, "You may recall that this frustration was sufficient so that I did not wish to attend later meetings."[45] These responses are interesting not only because they throw light on the conferences but also because they talk frankly about feelings. "Affect ran high," Savage recalled. In the transcripts, by contrast, emotions enter the discussion only as objects for scientific modeling. Almost never are they articulated as something the participants are experiencing. The contrast between the letters and the transcripts illuminates the scientific ethos that ruled at the meetings. Emotions were considered out of bounds for several reasons, all of which perhaps came down to the same reason. The framework of scientific inquiry had been constructed so as to ignore the observer.

Heinz von Foerster, in his letter to Fremont-Smith, saw the inclusion of the observer as the central issue of cybernetics.[46] He noted that at the beginning of the century, with the advent of relativity theory and the Uncertainty Principle, "a most enigmatic object was discovered which until then was carefully excluded from all scientific discourse: the 'observer.' 'Who is he?' was the question, indignantly asked by those who subscribed to a sour grape strategy, and seriously asked by those who felt that any science worth its name must include the subject that makes the observations at the first place." There were no precedents for this inclusion, he continued. "The whole methodology of a science that includes the observer had to be developed from scratch." He generously credited Fremont-Smith with the idea of bringing together people rather than disciplines and thus placing relationships at the center of the discussions (although the transcripts rarely acknowledge these relationships). He also commented that Fremont-Smith understood that including the observer would have to be an interdisciplinary task. In establishing the focus as "problems of communication," Fremont-Smith hoped the Macy group would see that the topic required an "intensive and comprehensive study of man." Thus the sciences were to be unified by an overarching framework that could simultaneously explain "man" and the people who studied "man." Cybernetics was to provide that framework.

In March 1976, two decades after the conferences had ended, Margaret Mead and Gregory Bateson were sitting with Stewart Brand at Bateson's kitchen table in a rare joint interview. Brand asked them about the Macy Conferences. They agreed that including the observer was one of the central problems raised by the cybernetic paradigm. Reaching for a scrap of paper, Bateson sketched a diagram (which Brand included in the published interview) of the communication system as it was envisioned before cybernetics. The drawing shows a black box with input, output, and feedback loops within the box. The space labeled "Engineer" remains outside the box. A second drawing represents Bateson's later understanding of cybernetics. Here the first black box, along with the names "Wiener, Bateson, Mead," is encapsulated within a larger box. In this drawing, the observers are included *within* the system rather than looking at it from the outside. The interview turned to a discussion of the dynamics that had prevailed at the Macy Conferences. Mead commented, "Kubie was a very important person at that point." She added: "McCulloch had a grand design in his mind. He got people into that conference, who he then kept from talking." Bateson continued, "Yes, he had a design for how the shape of the conversation would run over five years—what had to be said before what else could be said." When Brand asked what that design was, Bateson answered, "Who knows?" But Mead thought it was "more or less what happened."[47]

Brand wanted to know why cybernetics had run out of steam. "What happened?" he asked repeatedly. His sense of the situation is confirmed by correspondence exchanged between the transcript editors—Heinz von Foerster, Margaret Mead, and Hans Teuber—after the tenth conference in 1953. Fremont-Smith and McCulloch wanted the transcripts published, just as the transcripts for the previous four conferences had been published. But Teuber disagreed, noting that the discussions were too rambling and unfocused; if published, he said, they would be an embarrassment. Although he was the junior member of the editorial board, he stood his ground. He wrote to Fremont-Smith, sending a copy of the letter to McCulloch, that if the others decided to publish over his objections, he wanted his name removed from the list of editors.[48] As the junior member, he had the most to lose; the others already had established reputations. McCulloch must have written a stiff note in reply, for Teuber answered defensively. He insisted that the issue was not his reputation but the quality of the transcripts. "From your note, it is obvious that I sound stuffy to you and Walter. Do tell him that I wanted to get off the list of editors, not because I am worried about reputations, but simply because I can't do enough for this transcript to get it into any sort of shape. The transactions of this last meeting simply do not add to the earlier ones—they detract. Granted, there are a few sparks, but there is not enough of the old fire. I owed it to you and Frank Fremont-Smith to speak my mind on this matter."[49] Mead worked out a compromise. The three speakers would publish their talks as formal papers, and McCulloch's summary of all the conferences would be used as an introduction. No one thought of suggesting more conferences or more transcripts. It was the end of an era.

But not the end of reflexivity. Although a reflexive view of cybernetics failed to coalesce into a coherent theory during the Macy Conferences, Bateson did not want to let the idea go. He determined to go ahead on his own. He organized a conference in July 1968 to explore how the reflexive implications of cybernetics could provide the basis for a new epistemology, and he invited a group of scientists, social scientists, and humanists. Included were Warren McCulloch and Gordon Pask, both central players in cybernetics, along with Mary Catherine Bateson, known as Catherine (to her father as "Cap"), an anthropologist specializing in comparative religions.

Out of this week-long conference came Catherine's 1972 book, *Our Own Metaphor.*[50] Her account of this conference, in some ways a reflection of the Macy Conferences, contrasts sharply with the Macy transcripts. The best explanation for this difference, I think, is epistemological. Catherine assumes that *of course* the observer affects what is seen, so she takes care to tell her readers about her state of mind and situation at the time. She recounts, for example, finding out that she was pregnant in the months preceding the conference; how awed she felt by the life that, whether she consciously attended to it or not, continued to grow within her; and her devastation when the baby was born prematurely, lived for an afternoon, and died. Her grief was still fresh when she attended the conference, and it naturally colored, she feels, how she interpreted what she learned there.

The difference between her account and the Macy transcripts does not lie in the fact that one is technical and the other anecdotal. It is obviously important to Catherine to understand, as clearly as possible, what each presenter is saying, and she skillfully guides her reader through presentations fully as complex, technical, and detailed as any in the Macy transcripts. Rather, the difference lies in her attitude toward her material and her determination to include as much of the context as she can. She takes care to tell her readers not only what ideas were exchanged but also how the people looked and her interpretation of how they were feeling. In addition to the words exchanged, she includes appearance, body language, and emotional atmosphere. At the Macy Conferences, her mother, Margaret Mead, had repeatedly cautioned that the transcripts were a purely *verbal* record and therefore represented only a fraction of the communication taking place. Mead wanted a much fuller record that would include "posture, gesture, and intonation."[51] Two decades later, Catherine fulfilled that desire in her precisely crafted descriptions.

Here is Catherine's account of Warren McCulloch: "Warren had bright, fierce eyes and held his head dropped low between thin shoulders. He had white hair and a white beard and curious blend of glee and grief, of belligerence and gentleness" (*OOM*, pp. 23–24). When he gave a presentation, Catherine strained to follow his ideas and found it odd that he was not more responsive to the needs and situations of those who were listening. "More than anyone else present, Warren tended to use an uncompromisingly technical vocabulary, referring to scientists I knew nothing of and calling on unfamiliar mathematics and neurophysiology. As I listened I kept checking to see whether I was sorting out what each example was about, what kind of thing he was trying to say in this interdisciplinary context where not more than two or three people could follow the substance of most of his examples" (*OOM*, p. 65). In her contextualized account, McCulloch's fierce commitment to an "empirical epistemology" carries with it an obvious price—a tendency toward decontextualization that made him less than effective in communicating with this audience.

Catherine Bateson included in her prologue Gregory Bateson's document that set the agenda for the conference and laid out the problems it would explore. The influence of cybernetics as it had evolved during the Macy Conferences is apparent throughout. Equally clear are Gregory's revisions, critiques, and transformations of those concepts. He indicated that he wanted participants to consider "three cybernetic or homeostatic systems": the individual, the society, and the larger global ecosystem in which both are embedded. Although consciousness would be considered as "an important component in the *coupling* of these systems" (*OOM*, p. 13), epistemologically its role was limited. From an "enormously great plethora of mental events," it chooses a few on which to focus (*OOM*, p. 16). An important factor guiding this choice, he hypothesized, is "purpose." Problems arise when this purposeful selection is taken as the whole. "If consciousness has feedback upon the remainder of mind and if consciousness deals only with a skewed sample of the events of the total mind, then there must exist a systematic (i.e., non-random) difference between the conscious views of self and the world and the true nature of self and the world" (*OOM*, p. 16). Thus the emphasis on "purpose" so central to the Macy Conferences became here not an assumed orientation but a lens that consciousness wears and that distorts what it sees. Specifically, this lens obscures "the cybernetic nature of self and the world," an obfuscation that "*tends to be imperceptible to consciousness*" (*OOM*, p. 16).

Nowhere is the transformation that Gregory worked on the Macy Conferences clearer than in what he considers the "cybernetic nature" of world and self. For him, cybernetics is no longer the homeostatic model of the Macy Conferences (although echoes of this language still linger). Rather, it has become the reflexivity of the larger box that he would sketch a decade later at his kitchen table. Equally striking is the changed significance of separating a system from its surrounding context. For Bateson, decontextualization is not a necessary scientific move but a systematic distortion.

The inclination of the conscious mind toward purpose makes it focus on an arc of causally related events leading to a perceived goal. Obliterated or forgotten is the matrix in which these arcs are embedded. A truly cybernetic approach, for Bateson, concentrates on the couplings that bind the parts into interactive wholes.

The revisionist thrust of Gregory's view of cybernetics is apparent in a letter he wrote to Catherine in June 1977, a year after his interview with Stewart Brand. The letter begins with Gregory remarking on how rereading *Our Own Metaphor* vividly brought the conference back to his mind. Then Gregory lays out the gist of his new "cybernetic" epistemology. He starts from the premise that we never know the world as such. We know only what our sensory perceptions construct for us. In this sense, we know nothing about the world. But we know something, and what we know is the end result of the internal processes we use to construct our inner world. Thus we know ourselves as complex beings, including processes that extend below consciousness and beyond ourselves out into the world, through the inner world available to a consciousness that exists only because of those processes. "We are our epistemology" is Gregory's formulation.[52] Catherine's phrasing is similar: "Each person is his own central metaphor" (*OOM*, p. 285). In this view, the dualism between subject and object disappears, for the object as a thing in itself cannot exist for us. There is only the subjective, inner world. The world, as this "cybernetics" constructs it, is a monism. Nevertheless, it is not solipsistic, for Gregory believes that the microcosm of the inner world is functional within the larger ecosystem only because it is an appropriate metaphor for the macrocosm. In her concluding chapter, Catherine amplifies on this view by supposing that we can understand the complexity of the outer world only because our codes for constructing the inner world are similarly diverse and complex. In this sense, we are a metaphor not only for ourselves but also for the larger system in which we are embedded. This leads her into a subtly nuanced analysis of couplings between inner world and outer world, including the insight that because the worlds are coupled, they must in the last analysis be regarded as a single system.

For Gregory, McCulloch represents a Moses-like figure who could lead others to the brink of this new epistemology but was unable to enter into it himself. "His last speech makes a special sort of sense if you read it as spoken in that context," Gregory suggests.[53] Catherine uses McCulloch's speech to end her account of the conference, and the speech is worth quoting in detail. "I am by nature a warrior, and wars don't make sense anymore," McCulloch begins (*OOM*, p. 311). The recognition rings true. I think of the statement in his summary of the Macy Conferences: "Our most notable agreement is that we have learned to know one another a bit better, and to fight fair in our shirt sleeves."[54] For him, scientific debate was a form of agonistic conflict. He continues in his speech by recalling the nitty-gritty details of his experimental work, its difficulties and funny moments. Then his thoughts turn to human mortality. He is an old man; although he cannot know it now, within a year he will die. Earlier in the conference, he "snapped" (says Catherine): "I don't particularly like people. Never have. Man to my mind is about the nastiest, most destructive of all the animals. I don't see any reason, if he can evolve machines that can have more fun than he himself can, why they shouldn't take over, enslave us, quite happily. They might have a lot more fun. Invent better games than we ever did" (*OOM*, p. 226).

Now, at the penultimate moment of the conference, of Catherine's book that she will dedicate to him, and of his life, he confesses to mortal feelings. "'The difficulty is that we, who are not single-cell organisms, cannot simply divide and pass on our programs. We have to couple and there is behind this a second requirement.' Warren began to weep. 'We learn . . . that there's a utility in death because . . . the world goes on changing and we can't keep up with it. If I have any disciples, you can say this of every one of them, they think for themselves'" (*OOM*, p. 311).

If Gregory Bateson thought of himself as McCulloch's disciple, the epitaph that McCulloch wanted for himself is certainly true in Bateson's case, for he both learned from his mentor and went beyond him. Taking the cybernetic paradigm of McCulloch's "empirical epistemology" and making it into "our own metaphor," Bateson reintroduced the reflexive dimension that McCulloch had fought so hard to exorcise when it was associated with psychoanalysis. Yet Bateson's reinterpretation succeeded in articulating a version of reflexivity that did not depend on a psychoanalytic entanglement of conscious and unconscious meanings in scientific statements. Moreover, his epistemology gave an important role to objective constraints, for it insisted that only those constructions that were compatible with reality were conducive to long-term survival. And survival was very much the name of the game for Catherine and Gregory Bateson. The larger issues they wanted their conference to address included the increasing degradation of the environment. In looking for an epistemology that would proceed from a sense of the world's complexity, they did not give up the idea that some constructions are better than others.

Let me now anticipate connections between the path the Batesons followed and those paths traced in subsequent chapters. In breaking new conceptual ground, Gregory Bateson drew on a famous article on the frog's visual cortex. The article had been coauthored by several people from the Macy Conferences, including Warren McCulloch, Walter Pitts, and Jerome Lettvin; also listed as coauthor was a newcomer who did not attend the Macy Conferences, Humberto Maturana.[55] In using this article to develop "our own metaphor," Bateson went where no experimentalist could easily follow, for he made speculative leaps that would take decades of experimental work to confirm. He went into the inner world and turned it inside out, so to speak, so that the inner world became a metaphor for the outer world. Maturana was to follow a similar yet different path. He went into the inner world and insisted that it can't be turned inside out, that it is a metaphor for nothing other than its own creation of itself as a system. This is the theory of autopoiesis, which we will discuss in chapter 6. Maturana did not identify with cybernetics as much as Bateson did, and he did not generally use that term to describe his work. Nevertheless, his theory took up certain problems that were left hanging after the Macy Conferences ended. Like Bateson, Maturana found reflexivity more promising than homeostasis. Also like Bateson, he both appropriated concepts from the Macy context and changed them profoundly.

Janet Freud/Freed

Like Bateson, Mead, and Brand sitting at a kitchen table on that March morning in 1976, I am sitting at my kitchen table in March 1996. I'm looking at the pages on which their interview is published. I'm particularly intrigued by a photograph that Brand included, one evidently given to him by Mead or Bateson. It's a large picture, too large to include in one frame, so it stretches across two pages. The caption identifies the setting as the 1952 Macy Conference—the ninth, the conference with the last real Macy transcript, for the tenth volume (as noted above) was not a transcript but was instead formal papers. This was the conference of Kubie's last stand. The photograph shows a large group of men and one woman—Margaret Mead—sitting around cloth-covered tables pulled into a U-shape. A speaker stands at the mouth of the U; the caption identifies him as Yehoshua Bar-Hillel. But wait. That must mean the date is incorrect, since Bar-Hillel spoke at the tenth conference. He wasn't present at the ninth. So this photograph must have been taken in 1953, at the conference in which the conversation was so meandering and dilatory that it couldn't be published. I wonder where the caption came from. I imagine Bateson digging out the photograph and giving it to Brand while he and Mead clue Brand in on who was who as Brand scribbles down the names, probably while they are all still sitting at the kitchen table.

Now I notice that Mead isn't the only woman in the picture. Another woman sits with her back to the photographer, her arms extended, hands reaching out to a machine I can't quite see. The caption identifies her as "Janet Freud," but I know this can't be right either. She must be Janet Freed, listed in the published transcripts as "assistant to the conference program." I have seen her name in the typed transcripts of the editorial meetings that followed the later conferences, and I know more or less what she did.

She was responsible for turning these men's (and a couple of women's) words into type. She was the one who listened to the tape-recordings of the early conferences and strained to catch inaudible strange words. When she sent McCulloch the typed transcript of the second Macy conference, she plaintively wrote that she knew there were "many, many blank spaces" but that Dr. Fremont-Smith had ordered her and her staff to listen to the recordings only twice and to type what they heard.[56] Evidently, transcribing the tape-recordings was taking too much staff time, and Fremont-Smith did not want to waste his resources that way—his resources, her time.

The quirk of memory or handwriting that made Brand call her "Janet Freud" seems eerily appropriate, for this was the woman who, like Freud's patients, had no voice in the transcripts, although the transcripts have a voice that we can read only because of her. She was the one who presided over the physical transformations of signifiers as they went from tape-recording to transcript to revised copy to galley to book. Others—the editors Teuber, Mead, and von Foerster, the organizer Fremont-Smith, and the chairman McCulloch—worried about content—but her focus was the materiality of the processes that make sounds into words, marks into books. She did the best she could, but the transcription took much time and she had many other things to do. When she was told not to take time, the transcript had more ellipses than words, and she felt bad. What to do? She suggested to Fremont-Smith that he and McCulloch insist the speakers deliver drafts of their talks ahead of time.[57] Then she wouldn't have to strain to listen to tape-recordings that were noisy beyond endurance by today's standards. She wouldn't have to guess at unfamiliar words (the manuscripts of the transcripts are peppered with misspellings). She learned stenotypography (or perhaps arranged to hire someone else who knew it) so that the words could be transcribed directly into the machine. This, combined with the drafts of the presentations, allowed her to come up with reasonable transcripts of both presentation and discussion without driving herself crazy. At an editorial meeting, when others suggested that it was too much work to pressure the speakers to get their drafts into the office ahead of time, she spoke up. The drafts were essential. She defended the other woman who was lower on the totem pole than she was—her staff—and said that this woman could be expected to do only so much. She didn't say so, but surely she had herself in mind as well.

Janet Freed's role in the Macy group is teasingly hinted at in the transcripts to the 1949 Editors' Meeting. Fremont-Smith depended on her to keep him on track. He decided to make up a little booklet for the Macy Conference chairmen to supply them with guidelines, commenting, "It occurred to us, in fact, it was Miss Freed's suggestion . . . " Elsewhere, when he realized that he had "jumped around a good deal" and gotten off track, he referred to the list of topics that Freed had made up for him to follow.[58] When one of the men remarked that there were now thirteen Macy groups and wondered if his office was going "to be able to do it," Fremont-Smith must have looked at Freed, for he uttered a comment that, in this professional and overwhelmingly male meeting, comes across as almost shocking in its personal nature. "You write and get a lovely smile. Do you have anything else you want to say at this point?" "No," she replies, not elaborating. Nowhere else in the Macy transcripts, to my knowledge, does someone simply answer, "No." Perhaps she was embarrassed, or perhaps she simply felt her position made it inappropriate for her to say more.

Fremont-Smith's remark, faithfully preserved by the transcription technologies that Janet Freed oversaw, has a slightly odd phrasing, and I puzzle over it. She writes and gets a smile, as if she had to go somewhere to fetch it, as if it were produced elsewhere and transported back to her face. I feel I don't know where the smile comes from because Janet Freed effaces herself. Rarely do we see her directly; we glimpse her largely through her reflections in the speech of others. More than anyone else, she qualifies as the outside observer who watches a system that she constructs through the marks she makes on paper, although the system itself has a great deal of trouble including her within the names of those people who are authorized to speak and make meaning.

What are we to make of Janet F., this sign of the repressed, this Freudian slip of a female who, with a flick of a "u" (the U-shaped table at which she sits?), goes from Freed to Freud, Freud to Freed? Thinking of her, I am reminded of Dorothy Smith's suggestion that men of a certain class are prone to decontextualization and reification because they are in a position to command the labors of others.[59] "Take a letter, Miss Freed," he says. Miss

Freed comes in. She gets a lovely smile. The man speaks, and she writes on her stenography pad (or perhaps on her stenography typewriter). The man leaves. He has a plane to catch, a meeting to attend. When he returns, the letter is on his desk, awaiting his signature. From his point of view, what has happened? He speaks, giving commands or dictating words, and things happen. A woman comes in, marks are inscribed onto paper, letters appear, conferences are arranged, books are published. *Taken out of context,* his words fly, by themselves, into books. The full burden of the labor that makes these things happen is for him only an abstraction, a resource diverted from other possible uses, because he is not the one performing the labor.

Miss Freed has no such illusions. *Embedded in context,* she knows that words never make things happen by themselves—or rather, that the only things they can make happen are other abstractions, like getting married or opening meetings. They can't put marks onto paper. They can't get letters in the mail. They can't bring twenty-five people together at the right time and in the right place, at the Beekman Hotel in New York City, where white tablecloths and black chalkboards await them. For that, material and embodied processes must be used—processes that exist never in isolation but always *in contexts* where the relevant boundaries are permeable, negotiable, instantiated.

On a level beyond words, beyond theories and equations, in her body and her arms and her fingers and her aching back, Janet Freed knows that information is never disembodied, that messages don't flow by themselves, and that epistemology isn't a word floating through the thin, thin air until it is connected up with incorporating practices.