

# Ethics and Second-Order Cybernetics

Heinz von Foerster

1991

Ladies and Gentlemen:

I am touched by the generosity of the organizers of this conference who not only invited me to come to your glorious city of Paris, but also gave me the honor of opening the Plenary sessions with my presentation. And I am impressed by the ingenuity of the organizers who suggested to me the title of my presentation. They wanted me to address myself to “Ethics and Second-Order Cybernetics.” To be honest, I would have never dared to propose such an outrageous title, but I must say that I am delighted that this title was chosen for me.

Before I left California for Paris, others asked me full of envy, what am I going to do in Paris, what will I talk about? When I answered “I shall talk about Ethics and Second-Order Cybernetics” almost all of them looked at me in bewilderment and asked, “What is second-order cybernetics?” as if there were no questions about ethics. I am relieved when people ask me about second-order cybernetics and not about ethics, because it is much easier to talk about second-order cybernetics than it is to talk about ethics. In fact it is impossible to talk about ethics. But let me explain that later, and let me now say a few words about cybernetics, and of course, the cybernetics of cybernetics, or second-order cybernetics.

As you all know, cybernetics arises when effectors (say, a motor, an engine, our muscles, etc.) are connected to a sensory organ which in turn acts with its signals upon the effectors. It is this circular organization which sets cybernetic systems apart from others that are not so organized. Here is Norbert Wiener, who re-introduced the term “Cybernetics” into scientific discourse. He observed, “The behavior of such systems may be interpreted as directed toward the attainment of a goal.” That is, it looks as if these systems pursued a purpose!

That sounds very bizarre indeed! But let me give you other paraphrases of what cybernetics is all about by invoking the spirit of women and men who rightly could be considered the mamas and papas of cybernetic thought and action. First there is Margaret Mead, whose name I am sure is familiar to all of you. In an address to the American Society of Cybernetics she remarked:

As an anthropologist, I have been interested in the effects that the theories of Cybernetics have within our society. I am not referring to computers or to the electronic revolution as a whole, or to the end of dependence on script for knowledge, or to the way that dress has succeeded the mimeographing machine as a form of communication among the dissenting young. Let me repeat that, I am *not* referring to the way that *dress* has succeeded the mimeographing machine as a form of communication among the dissenting young.

And she then continues:

I specifically want to consider the significance of the set of cross-disciplinary ideas which we first called “feed-back” and then called “teleological mechanisms” and then called it “cybernetics,” a form of cross-disciplinary thought which made it possible for members of many disciplines to communicate with each other easily in a language which all could understand.

And here is the voice of her third husband, the epistemologist, anthropologist, cybernetician, and as some say, the papa of family therapy, Gregory Bateson, “Cybernetics is a branch of mathematics dealing with problems of control, recursiveness and information.

And here is the organizational philosopher and managerial wizard Stafford Beer, “Cybernetics is the science of effective organization.”

And finally, here the poetic reflection of “Mister Cybernetics,” as we fondly call him, the Cybernetician’s cybernetician; Gordon Pask, “Cybernetics is the science of defensible metaphors.”

It seems that cybernetics is many different things to many different people. But this is because of the richness of its conceptual base; and I believe that this is very good, otherwise cybernetics would become a somewhat boring exercise. However, all of those perspectives arise from one central theme; that of circularity. When, perhaps a half century ago, the fecundity of this concept was seen, it was sheer euphoria to philosophize, epistemologize, and theorize about its unifying power and its consequences and ramification on various fields. While this was going on, something strange evolved among the philosophers, the epistemologists and, the theoreticians. They began to see themselves more and more as being included in a larger circularity; maybe within the circularity of their family; or that of their society and culture; or even being included in a circularity of cosmic proportions!

What appears to us today as being most natural to see and think, was then not only difficult to see, but wasn’t even allowed to be thought. Why? Because it would violate the basic principle of scientific discourse which demands the separation of the observer from the observed. It is the principle of objectivity. The properties of the observer shall not enter the description of his observations.

I present this principle here, in its most brutal form, to demonstrate its non-sensicality. If the properties of the observer (namely to observe and describe) are eliminated, there is nothing left; no observation, no description. However, there was a justification for adhering to this principle, and this justification was fear; fear that paradoxes would arise when the observers were allowed to enter the universe of their observations. And you know the threat of paradoxes. To steal their way into a theory is like having the cloven-hoofed foot of the devil stuck in the door of orthodoxy.

Clearly when cyberneticians were thinking of partnership in the circularity of observing and communicating, they were entering into a forbidden land. In the general case of circular closure, A implies B; B implies C; and (Oh, horror!) C implies A! Or in the reflexive case, A implies B, and (Oh, shock!) B implies A! And now the devil’s cloven-hoof in its purest form, the form of self-reference; A implies A (Outrage!)

I would like to invite you now to join me in a land where it is not forbidden; rather, where one is encouraged to speak about oneself. What else can one do anyway? This turn from looking at things “out there” to looking at “looking itself,” arose I think, from significant advances in neurophysiology and neuropsychiatry. It appeared that one could now dare to ask the question of how the brain works. One could dare to write a theory of the brain.

It may be argued that over the centuries since Aristotle, physicians and philosophers again and again developed theories of the brain. So, what’s new of today’s cyberneticians? What is new is the profound insight that a brain is required to write a theory of a brain. From this follows that a theory of the brain, that has any aspirations for completeness, has to account for the writing of this theory. And even more fascinating, the writer of this theory has to account for her or himself. Translated into the domain of cybernetics; the cybernetician, by entering his own domain, has to account for his or her own activity. Cybernetics then becomes cybernetics of cybernetics, or *second-order cybernetics*.

Ladies and Gentlemen, this perception represents a fundamental change, not only in the way we conduct science, but also how we perceive teaching, learning, the therapeutic process, organizational management, and so on and so forth; and I would say, of how we perceive relationships in our daily life. One may see this fundamental epistemological change if one first considers oneself to be an independent observer who watches the world go by; as opposed to a person who considers oneself to be a participant actor in the drama of mutual interaction of the give and take in the circularity of human relations.

In the case of the first example, as a result of my independence, I can tell others how to think and act, “Thou shalt . . .” “Thou shalt not . . .” This is the origin of moral codes. In the case of the second example, because of my interdependence, I can only tell myself how to think and act, “I shall . . .” “I shall not . . .” This is the origin of ethics. This was the easy part of my presentation. Now comes the difficult part. I am supposed to talk about ethics. How to go about this? Where to begin?

In my search for a beginning I came across the lovely poem by Yveline Rey and Bernard Prieur that embellishes the first page of our program. Let me read to you the first few lines:

“Vous avez dit Ethique?”  
Déjà le murmure s’amplifie en rumeur.  
Soudain les roses ne montrent plus des épines.  
Sans doute le sujet est-il brûlant.  
Il est aussi d’actualité.

Let me begin with epines—with the thorns—and I hope, a rose will emerge. The thorns I begin with are Ludwig Wittgenstein’s reflections upon ethics in his *Tractatus Logico-Philosophicus*. If I were to provide a title for this tractatus, I would call it *Tractatus Ethico-Philosophicus*. However, I am not going to defend this choice, I rather tell you what prompts me to refer to Wittgenstein’s reflections in order to present my own.

I’m referring to point Number 6 in his *Tractatus* where he discusses the general form of propositions. Near the end of this discussion he turns to the problem of values in the world and their expression in propositions. In his famous point Number 6.421 he comes to a conclusion which I will read to you in the original German, “Es ist Klar, dass sich Ethik nicht aussprechen laesst.” I wish I knew a French translation. I only know two English translations which are both incorrect. Therefore, I will present *my* translation into English, with my conviction that the simultaneous translators will do a superb job of presenting Wittgenstein’s point in French. Here is my English version of 6.421, “It is clear that ethics cannot be articulated.”

Now you understand why earlier I said, “My beginning will be thorns.” Here is an International Congress on Ethics, and the first speaker says something to the effect that it is impossible to speak about ethics! But please be patient for a moment. I quoted Wittgenstein’s thesis in isolation. Therefore it is not yet clear what he wanted to say.

Fortunately, the next point 6.422, which I will read in a moment, provides a larger context for 6.421. To prepare for what you are about to hear, you should remember that Wittgenstein was a Viennese. So am I. Therefore there is a kind of underground understanding which I sense you Parisians will share with us Viennese. Let me try to explain. Here now is point 6.422 in the English translation by Pears and McGuinness; “When an ethical law of the form “Thou shalt . . .” is laid down, one’s first thought is, ‘And what if I do not do it?’” When I first read this, my thought was that not everybody will share Wittgenstein’s view. I think that this reflects his cultural background.

Let me continue with Wittgenstein, “It is clear however, that ethics has nothing to do with punishment and reward in the usual sense of the terms. Nevertheless, there must indeed be some kind of ethical reward and punishment, but they must reside in the action itself.”

They must reside in the action itself! You may remember, we came across such self-referential notions earlier with the example, “A implies A” and its recursive relatives of second-order cybernetics. Can we take a hint from these comments for how to go about reflecting about ethics, and at the same time adhere to Wittgenstein’s criterion? I think we can. I myself try to adhere to the following rule; to master the use of my language so that ethics is implicit in any discourse I may have. (e.g., in science, philosophy, epistemology, therapy, etc.)

What do I mean by that? By that I mean to let language and action ride on an underground river of ethics, and to make sure that one is not thrown off. This insures that ethics does not become explicit and that language does not degenerate into moralizations. How can one accomplish this? How can one hide ethics from all eyes and still let her determine language and action? Fortunately, ethics has two sisters who allow her to remain unseen. They create for us a visible framework; a tangible tissue within which and upon which we may weave the goblins of our life. And who are these two sisters? One is Metaphysics, the other is Dialogics.

My job now is to talk about these two ladies, and how they manage to allow ethics to become manifest without becoming explicit.

## Metaphysics

Let me first talk about Metaphysics. In order to let you see at once the delightful ambiguity that surrounds her, let me quote from a superb article, “The Nature of Metaphysics” by the British scholar W.H. Walsh. He begins his article with the following sentence, “Almost everything in metaphysics is controversial, and it is therefore not surprising that there is little agreement among those who call themselves metaphysicians about what precisely it is they are attempting.”

Today, when I invoke Metaphysics, I do not seek agreement with anybody else about her nature. This is because I want to say precisely what it is when we become metaphysicians, whether or not we call ourselves metaphysicians. I say that we become a metaphysician any time we decide upon in principle undecidable questions. For instance, here is a decidable question, “Is the number 3,396,714 divisible by 2?” It will take you less than two seconds to decide that indeed this number is divisible by two. The interesting thing here is that it will take you exactly the same short time to decide if the number has not 7, but 7000 or 7 million digits. I could of course invent questions that are slightly more difficult; for instance, “Is 3,396,714 divisible by three?”, or even more difficult ones. But there are also problems that are extraordinarily difficult to decide, some of them having been posed more than 200 years ago and remain unanswered.

Think of Fermat’s “Last Theorem” to which the most brilliant heads have put their brilliant minds and have not yet come up with an answer. Or think of Goldbach’s “Conjecture” which sounds so simple that it seems a proof cannot be too far away, “All even numbers can be composed as the sum of two primes.” For example, 12 is the sum of the two prime numbers 5 and 7; or  $20 = 17 + 3$ ; or  $24 = 13 + 11$ , and so on and so forth. So far, no counterexample to Goldbach’s conjecture has been found. And even if all further tests would not refute Goldbach, it still would remain a conjecture until a sequence of mathematical steps is found that decides in favor of his good sense of numbers. There is a justification for not giving up and for continuing the search for finding a sequence of steps that would prove Goldbach. It is that the problem is posed in a framework of logico-mathematical relations which guarantees that one can climb from any node of this complex crystal of connections to any other node.

One of the most remarkable examples of such a crystal of thought is Bertrand Russell’s and Alfred North Whitehead’s monumental *Principia Mathematica* which they wrote over a 10 year period between 1900 and 1910. This 3 volume *magnum opus* of more than 1500 pages was to establish once and for all a conceptual machinery for flawless deductions. A conceptual machinery that would contain no ambiguities, no contradictions and no undecidables.

Nevertheless, in 1931, Kurt Gödel, then 25 years of age, published an article whose significance goes far beyond the circle of logicians and mathematicians. The title of this article I will give you now in English, “On formally undecidable propositions in the *Principia Mathematica* and related systems.” What Gödel does in his paper is to demonstrate that logical systems, even those so carefully constructed by Russell and Whitehead, are not immune to undecidables sneaking in.

However, we do not need to go to Russell and Whitehead, Gödel, or any other giants to learn about in principle undecidable questions. We can easily find them all around. For instance, the question about the origin of the universe is one of those in principle undecidable questions. Nobody was there to watch it. Moreover, this is apparent by the many different answers that are given to this question. Some say it was a single act of creation some 4 or 5,000 years ago. Others say there was never a beginning and that there will never be an end; because the universe is a system in perpetual equilibrium. Then there are those who claim that approximately 10 or 20 billion years ago the universe came into being with a “Big Bang” whose remnants one is able to hear over large radio antennas. But I am most inclined to trust Chuang Tse’s report, because he is the oldest and was therefore the closest to the event. He says:

Heaven does nothing, this nothing-doing is dignity;  
Earth does nothing, this nothing-doing is rest;  
From the union of these two nothing-doings arise all action  
And all things are brought forth.

I could go on and on with other examples, because I have not yet told you what the Burmese, the Australians, the Eskimos, the Bushmen, the Ibos, etc., would tell you about their origins. In other words, tell me how the universe came about, and I will tell you who you are.

I hope that I have made the distinction between decidable and, in principle, undecidable questions sufficiently clear so that I may present the following proposition which I call the “metaphysical postulate:”

Only those questions that are in principle undecidable, *we* can decide.

Why? Simply because the decidable questions are already decided by the choice of the framework in which they are asked, and by the choice of the rules used to connect what we label “the question” with what we take for an “answer.” In some cases it may go fast, in others it may take a long, long time. But ultimately we arrive after a long sequence of compelling logical steps at an irrefutable answer; a definite “yes,” or a definite “no.”

But we are under no compulsion, not even under that of logic, when we decide on in principle undecidable questions. There is no external necessity that forces us to answer such questions one way or another. We are free! The compliment to necessity is not chance, it is choice! *We can choose who we wish to become when we have decided on an in principle undecidable question.*

That is the good news, as American journalists would say, now comes the bad news. With this freedom of choice we are now responsible for the choice we make. For some, this freedom of choice is a gift from heaven. For others such responsibility is an unbearable burden. How can one escape it? How can one avoid it? How can one pass it on to somebody else?

With much ingenuity and imagination, mechanisms have been contrived by which one could bypass this awesome burden. Through hierarchies, entire institutions have been built where it is impossible to localize responsibility. Everyone in such a system can say, “I was told to do ‘X.’” On the political stage, we hear more and more the phrase of Pontius Pilate, “I have no choice but ‘X.’” In other words, “Don’t hold me responsible for ‘X.’ Blame someone else.” This phrase apparently replaces, “Among the many choices I had, I decided on ‘X.’”

I mentioned objectivity before, and I mention it here again as a popular device for avoiding responsibility. As you may remember, objectivity requires that the properties of the observer be left out of any descriptions of his observations. With the essence of observing (namely the processes of cognition) having been removed, the observer is reduced to a copying machine with the notion of responsibility successfully juggled away.

Objectivity, Pontius Pilate, hierarchies, and other devices are all derivations of a choice between a pair of in principle undecidable questions which are, “Am I *apart from* the universe?” Meaning whenever I *look*, I’m looking as if through a peephole upon an unfolding universe; or, “Am I *part of* the universe?” Meaning whenever I *act*, I’m changing myself and the universe as well.

Whenever I reflect on these two alternatives, I’m surprised by the depth of the abyss that separates the two fundamentally different worlds that can be created by such a choice. That is to see myself as a citizen of an independent universe, whose regulations, rules and customs I may eventually discover; or to see myself as a participant in a conspiracy, whose customs, rules, and regulations we are now inventing.

Whenever I speak to those who have made their decision to be either discoverers or inventors, I’m impressed by the fact that neither of them realizes that they have ever made that decision. Moreover, when challenged to justify their position, a conceptual framework is constructed which itself turns out to be the result of a decision upon an in principle undecidable question.

It seems as though I’m telling you a detective story while keeping quiet about who is the good guy and who is the bad guy; or who is sane and who is insane; or who is right and who is wrong. Since these are in principle undecidable questions, it is for each of us to decide, and then take responsibility for. There is a murderer. I submit that it is

unknowable whether he is or was insane. The only thing we *know* is what I say, what you say, or what the expert says he is. And what I say, what you say, and what the expert says about his sanity or insanity is my, is your, and is the expert's responsibility. Again, the point here is not the question "Who's right and who's wrong?" This is an in principle undecidable question. The point here freedom; freedom of choice. It is José Ortega y Gasset's point:

Man does not have a nature, but a history. Man is nothing but a drama. His life is something that has to be chosen, made up as he goes along. And a human consists in that choice and invention. Each human being is the novelist of himself, and though he may choose between being an original writer and a plagiarist, he cannot escape choosing. He is condemned to be free.

You may have become suspicious of me qualifying all questions as being in principle undecidable questions. This is by no means the case. I was once asked how the inhabitants of such different worlds as I sketched before, (the inhabitants of the world they discover, and the inhabitants of a world they invent) can ever live together. Answering that is not a problem. The discoverers will most likely become astronomers, physicists and engineers; the inventors family therapists, poets, and biologists. And living together won't be a problem either, as long as the discoverers discover inventors, and the inventors invent discoverers. Should difficulties develop, fortunately we have this full house of family therapists who may help to bring sanity to the human family.

I have a dear friend who grew up in Marakesh. The house of his family stood on the street that divides the Jewish and the Arabic quarters. As a boy, he played with all the others, listened to what they thought and said, and learned of their fundamentally different views. When I asked him once who was right he said, "They are both right."

"But this cannot be," I argued from an Aristotelian platform, "Only one of them can have the truth!"

"The problem is not truth," he answered, "The problem is trust."

I understood. The problem is understanding. The problem is understanding understanding! The problem is making decisions upon in principle undecidable questions.

At that point Metaphysics appeared and asked her younger sister Ethics, "What would you recommend I bring back to my proteges, the metaphysicians, regardless of whether or not they refer to themselves as such?" Ethics answered, "Tell them they should always try to act so as to increase the number of choices. Yes, increase the number of choices!"

## Dialogics

Now I would like to turn to Ethics' sister, Dialogics. What are the means at her disposal to insure that Ethics can manifest herself without becoming explicit? You may already have guessed that it is, of course, language. I am not referring here in the sense of the noises produced by pushing air past our vocal cords; or language in the sense of grammar, syntax, semantics, semiotics; nor the machinery of phrases, verb phrases, noun phrases, deep structure, etc. When I refer here to language, I refer to language the "dance." Similar to when we say "It takes two to Tango," I am saying, "It takes two to language."

When it comes to the dance of language, you the family therapists are of course the masters, while I can only speak as an amateur. Since "amateur", comes from "amour," you'll know at once that I love to dance this dance. In fact, what little I know of this dance I learned from you. My first lesson came when I was invited to sit in an observation room and observe through, the one way mirror a therapeutic session in progress with a family of four. For a moment my colleagues had to leave, and I was by myself. I was curious as to what I would see when I couldn't hear what was said, so I turned off the sound.

I recommend that you perform this experiment yourself. Perhaps you will be as fascinated as I was. What I saw then, the silent pantomime, the parting and closing of lips, the body movements, the boy who only once stopped biting his nails . . . what I saw then were the dance steps of language, the dance steps alone, without the disturbing effects of the music. Later I heard from the therapist that this session was very successful indeed. I thought, what

magic must sit in the noises these people produced by pushing air past their vocal cords and by parting and closing their lips. Therapy! What magic indeed! And to think that the only medicine at your disposal are the dance steps of language and its accompanying music. Language! What magic indeed!

It is left to the naive to believe that magic can be explained. Magic cannot be explained. Magic can only be practiced, as you all well know. Reflecting on the magic of language is similar to reflecting upon a theory of the brain. As much as one needs a brain to reflect upon a theory of the brain, one needs the magic of language to reflect upon the magic of language. It is the magic of those notions that they need themselves to come into being. *They are of second-order*. It is also the way language protects itself against explanation by always speaking about itself

There is a word for language, namely “language.” There is a word for word, namely “word.” If you don’t know what word means, you can look it up in a dictionary. I did that. I found it to be an “utterance.” I asked myself, “What is an utterance?” I looked it up in the dictionary. The dictionary said that it means “to express through words.” So here we are back where we started. Circularity; A implies A.

But this is not the only way language protects itself against explanation. In order to confuse her explorer she always runs on two different tracks. If you chase language up one track, she jumps to the other. If you follow her there, she is back on the first. What are these two tracks? One track is the track of appearance. It runs through a land that appears stretched out before us; the land we are looking at as though through a peephole. The other track is the track of function. It runs through the land that is as much a part of us as we are a part of it; the land that functions like an extension of our body.

When language is on the track of appearance it is a monologue. There are noises produced by pushing air past vocal cords. There are the words, the grammar, the syntax, the well formed sentences. Along with these noises goes the denotative pointing. Point to a table, make the noise “table”; point to a chair, make the noise “chair.”

Sometimes it does not work. Margaret Mead quickly learned the colloquial language of many tribes by pointing to things and waiting for the appropriate noises. She told me that once she came to a particular tribe, pointed to different things, but always got the same noises, “chumulu.” A primitive language she thought, only one word! Later she learned that “chumulu” means “pointing with finger.”

When language switches to the track of function it is dialogic. There are, of course, these noises; some of them may sound like “table,” others like “chair.” But there need not be any tables or chairs because nobody is pointing at tables or chairs. These noises are invitations to the other to make some dance steps together. The noises “table” and “chair” bring to resonance those strings in the mind of the other which, when brought to vibration, would produce noises like “table” and “chair.” Language in its function is connotative.

In its appearance, language is descriptive. When you tell your story, you tell it as it was; the magnificent ship, the ocean, the big sky, and the flirt you had that made the whole trip a delight. But for whom do you tell it? That’s the wrong question. The right question is; with whom are you going to dance your story, so that your partner will float with you over the decks of your ship, will smell the salt of the ocean, will let the soul expand over the sky? And there will be a flash of jealousy when you come to the point of your flirt.

In its function, language is constructive because nobody knows the source of your story. Nobody knows, nor ever will know how it was, because “as it was” is gone forever.

You remember René Descartes as he was sitting in his study, not only doubting that he was sitting in his study, but also doubting his existence. He asked himself, “Am I, or am I not?” “Am I, or am I not?” He answered this rhetorical question with the solipsistic monologue, “Je pense, donc je suis” or in the famous Latin version, “Cogito ergo sum.” As Descartes knew very well, this is language in its appearance, otherwise he would not have quickly published his insight for the benefit of others in his “Discourse de la méthode.” Since he understood the function of language as well, in all fairness he should have exclaimed, “Je pense, donc nous sommes”, “Cogito ergo sumus” or, “I think, therefore *we* are.”

In its appearance, the language I speak is *my* language. It makes me aware of myself. This is the root of *consciousness*. In its function, my language reaches out for the other. This is the root of *conscience*. And this is where Ethics

invisibly manifests itself through dialogue. Permit me to read to you what Martin Buber says in the last few lines of his book *Das Problem des Menschen*:

Contemplate the human with the human, and you will see the dynamic duality, the essence together. Here is the giving and the receiving, here is the aggressive and the defensive power, here the quality of searching and of responding, always both in one, mutually complementing in alternating action, demonstrating together what it is; human. Now you can turn to the single one and you can recognize him as human for his potential of relating. We may come closer to answering the question, “What is human?” when we come to understand him as the being in whose dialogic, in his mutually present two-getherness, the encounter of the one with the other is realized and recognized at all times.

Since I cannot add anything to Buber’s words, this is all I can say about ethics, and about second-order cybernetics. Thank you very much.

## **Yveline Rey: Interview with Heinz von Foerster**

**Yveline:** The first time I heard your name mentioned, it was accompanied by the term “cybernetician.” How does one become a cybernetician? Why this choice at the beginning? What were the influential steps throughout the course of your life?

**Heinz:** Yes. How does one become a cybernetician? Or, perhaps you want me to tell you how *I* became a cybernetician.

You may remember the point I made in my address; that we all are metaphysicians, whether we call ourselves such, whenever we decide upon in principle undecidable questions. To answer your question, I could also say we are all cyberneticians (whether or not we call ourselves such) whenever we justify our actions without using the words “because of . . .” or “à cause de . . .” but with the phrase in English “in order to . . .” which in French is much more Aristotelian, “à fin de . . .”

**Yveline:** Why Aristotelian?

**Heinz:** In his *Metaphysics*, Aristotle distinguished four different kinds of causes or, as I would say, four different excuses; two of which have temporal character, “causa efficientis” and “causa finalis.” Physicists love the former, where causes in the past determine the effects in the present: “*Because* she did turn the switch, the lights go on now.” Psychologists prefer the latter: “*In order to* have the lights on, she turns the switch now.” Causes in the future, “to have the room lit,” determine actions in the present, “turn the switch now.”

**Yveline:** Very interesting, but where does cybernetics come in?

**Heinz:** Physicists explore the connection between the positions of the switch, making or breaking contact, and the electrical processes that heat the wires in the lamp to temperatures that are high enough to radiate electromagnetic waves in the visible spectrum, etc., etc. Cyberneticians explore the connection between the little girl’s wish to enter a lit as opposed to a dark room, as well as the senso-motoric processes and the emerging eye-hand correlation that bring her hand along an unpredictable path, but with a predictable outcome, closer and closer to the switch which she then turns in the right direction, etc. If one were to watch this girl, one might be tempted to say as did Norbert Wiener, “. . . her behavior may be interpreted as directed to the attainment of a goal.” In the early cybernetic literature you will find again and again reference to the notion of “goal,” “purpose,” “end,” etc. Since the Greek word for “end” is “telos,” our pre-cyberneticians used “teleology” for identifying their activity.

**Yveline:** But, Heinz, you said before that we are all cyberneticians, whether or not we call ourselves such, but when I go to turn on a light switch I am not “exploring the senso-motoric connections. . .” et cetera. I just go and turn on the switch. Where is the cybernetician?

**Heinz:** (Laughing) This is one more reason why I love women! You look through all the scientific verbal haze and go straight to the essential points. Now ... Hmm ... What can I say?

I think I can extricate myself from this dilemma by inventing a new category of cybernetics: “Zero-order Cybernetics.” I suggest we have a case of zero-order cybernetics when activity becomes structured; when “behavior” emerges, but one doesn’t reflect upon the “why” and the “how” of this behavior. One just acts. This is when cybernetics is implicit.

**Yveline:** I see. But what is now “First-order cybernetics?”

**Heinz:** This is when one reflects upon one’s behavior, upon the “how” and the “why.” Then cybernetics becomes explicit, and one develops notions like “feedback,” “amount of information,” “circularity,” “recursion,” “control” “homeostasis,” “dynamic stability,” “dynamic instability or chaos,” “fixed points,” “attractors,” “equi-finality,” “purpose,” “goal,” etc., etc. In other words, one arrives at the whole conceptual machinery of “early” cybernetics, first-order cybernetics; or as I would say, the cybernetics of observed systems.

**Yveline:** Let me come back to my first question. How did you come upon cybernetics?

**Heinz:** Very simple. Cybernetics came upon me; because my English vocabulary was at most 25 words.

**Yveline:** This makes no sense, dear Heinz. You’ll have to explain that a bit better.

**Heinz:** Okay. Then we have to go back to a time when you, dear Yveline, were not yet born. We have to go back to the year 1948, when parts of Austria were still occupied by Russian troops, and the world was slowly recovering from the wounds of the war. In November of that year, in Cambridge, Massachusetts, Norbert Wiener published a book entitled *Cybernetics*, with the subtitle *Communication and Control in the Animal and the Machine*. Also that November, Heinz von Foerster in Vienna, Austria, published a book entitled *Das Gedächtnis* [The memory] with the subtitle *Eine quantenphysikalische Untersuchung*, [An investigation in quantum physics]. I am originally a physicist, and what I tried to do in this investigation was to connect observations in experimental psychology and neurophysiology with the physics of the large (biological) molecules. I think that I didn’t do a bad job of it.

Now I have to switch to another track. My wife’s dearest friend, Ilse, had escaped from Germany when Hitler came into power. By 1948 she was well established in New York and she invited me to come to the United States in the hope that I could establish a beachhead in order to make it easier for the rest of my family to follow. In February of 1949 I crossed a very stormy Atlantic on the *Queen Mary*. Since I don’t get seasick, (most of the other passengers were) I always had 6 waiters serving me in an empty dining room.

A few days after my arrival in New York, one of America’s leading neuropsychiatrist, Warren McCulloch (who, by an amazing combination of miraculous circumstances, had gotten hold of my book) invited me to present my theory of memory at a meeting in New York that was to take place a few days later. He also recommended that I find a book entitled *Cybernetics* in order to prepare myself a bit for this meeting. I did that, and with the little English at my disposal at that time, I tried hard to understand some of its basic points.

Somewhat ill prepared in concepts and language, I came to this meeting whose title was more or less an enigma as well: “Circular Causal and Feedback Mechanisms in Biological and Social Systems.” To my surprise, it was a small meeting of about 20 participants, but to my even greater surprise, this was an assembly of the *crème de la crème* of American scientists. There was, of course, Warren McCulloch who was chairman of the conference, and whose works in 4 volumes have recently been published. There was Norbert Wiener himself, of whom a lovely biography by P. Masani appeared last year. There was John von Neuman, the man who started the computer revolution. Then there were Gregory Bateson and his wife Margaret Mead, or should I say Margaret Mead and her husband Gregory Bateson, who brought to anthropology wisdom, profundity and humor; both in different ways.

These are but a few, whose names I believe would be familiar to my European friends. I don’t know who invented the notion of “interdisciplinarity,” but this meeting was its manifestation. If you were to begin with

Anthropology in an alphabetical list of academic professions, and end with Zoology, my guess would be that almost every one of these disciplines had a representative present.

I was called upon relatively early to present my story, and I wrestled valiantly with my 20 English words to make myself understood. The whole thing would have turned into a catastrophe if it weren't for the presence of Gerhard von Bonin, Heinrich Klüver and others who spoke fluently German and who rescued me by translating some of my arguments.

That evening, the group had a business meeting. Before it was over, I was invited to come in. "Heinz," began the chairman, "we listened to your molecular theory of memory, and your theory agrees with many observations which other theories cannot account for. *What* you had to say was very interesting. However, *how* you said it was abominable! Because we want you to learn English fast, we have decided to appoint you to be the editor of the transactions of this conference."

I was of course speechless. How could I edit articles by such superb writers as Wiener, Mead, Bateson, etc.? How could I organize material of which I, at best, understood only half? But, I thought "Why not try?". So I accepted the appointment. I immediately proposed that, "Since the title of this conference is so long, it is hard to remember, and for me, hard to pronounce; 'circular-causal-and-feedback-mechanisms...' I propose to call this conference 'Cybernetics.'"

Everybody looked at Norbert Wiener, who sat next to me, and applauded in his honor and in acceptance of my proposal. Deeply touched by the recognition of his peers, tears came to his eyes, and he left the room to hide his emotions.

The sponsor of this, and four more conferences on this topic, was the Josiah Macy Jr. Foundation of New York, who asked me to edit each of the 5 volumes. Since all of that took place in the remote past, aficionados of cybernetics refer to these books as the "legendary Macy meetings on cybernetics."

Here ends, dear Yveline, my story of how cybernetics came upon me.

**Yveline:** Throughout the course of the conference, in the conference rooms as well as the corridors of the Cité de la Villette, there was much discussion about first-order cybernetics and second-order cybernetics; mostly to put them opposite each other. For instance, "But you see my dear, in my view this is from first-order cybernetics..." or, "I tell you, one really feels the difference; this time we are in the second-order cybernetics." Would you attempt to clarify for the people here, what are the fundamental distinctions for first-order and second-order cybernetics? Which change in direction or observation signify second-order cybernetics? Or to paraphrase G. Spencer-Brown, whom you like to cite, "Design me a resemblance!" or, "Design me a distinction!"

**Heinz:** Let me draw the distinction for you. You followed me when I moved from zero-order to first-order cybernetics. What did I do? I let the underlying circularity of processes of emergence, of manifestation, of structurization, of organization, etc., become explicit. By that I mean that we now *reflect* about these circular processes which generate structure, order, behavior, etc., in those things we observe. Now Yveline, you can easily guess how to move from first-order to second-order cybernetics.

**Yveline:** I think so. Let me try. In second-order you reflect upon your reflections.

**Heinz:** Of course!

**Yveline:** And now, can I go on to third-order cybernetics?

**Heinz:** Yes, you could. But it would not create anything new, because by ascending into "second-order," as Aristotle would say, one has stepped into the circle that closes upon itself. One has stepped into the domain of concepts that apply to themselves.

Y Do you mean to say that a second-order cybernetics is a cybernetics of cybernetics?

**Heinz:** Yes, precisely!

**Yveline:** Can you give me other examples?

**Heinz:** Yes of course. For instance, compare a typical first-order cybernetics concept such as “purpose,” (as being the equivalent of “why”) with a second-order question, “What is the purpose of ‘purpose’?” (asking why the notion of “purpose” is used in the first place; i.e. how does it influence discourse, explanations, argumentations, etc.?)

One nice feature of this notion is that it relieves one of the need to account for the way things are done which are intended. Every time I tie my shoelaces, or you slip into your pumps, we do it differently. We do it in thousands of unpredictable variations, but the outcome is predictable; my shoelaces are tied, your shoes are on your feet.

On the other hand, it is quite impossible for a physicist to invent the “Laws of Nature” with which to compute our behavior from the initial conditions of my united shoelaces or your pumps in your wardrobe; that is to compute the paths, the “trajectories” and the movements that our bodies and our limbs are taking, which tie my laces or put shoes on your feet. The physicist’s “causa efficientis” is impotent. But the cyberneticist’s “causa finalis” does it all. If the intentions are clear, (independent of the initial conditions) the sensorimotor loops will adjust and readjust our movements until my laces are tied; your shoes are on your feet.

**Yveline:** Thank you. I feel much better with my shoes on. I see now the purpose of using the notion of purpose. One does not need to know *how* to get there; one needs only to know the *there*. This is a very nice feature indeed! Is there a bad feature too?

**Heinz:** Yes there is. The ugly feature of the notions of “purpose,” “goal,” “end,” is that they can be used to justify the specific ways of getting there; “The end justified the means.” And as we know now, the means can be very ugly indeed. The question should be, “Do the means justify the end?”

**Yveline:** If we would remember to ask the question this way, the world could be a very different place. But now Heinz, to use your language, tell me how did second-order cybernetics “come upon” you?

**Heinz:** Through a woman, of course. It was Margaret Mead. You remember the quote I cited in my address? It came from a speech she gave, I think in 1968. Since she rarely uses titles for her talks and almost never reads from a script, I sent her the transcript from a recording asking for her corrections and a title. There was no reply. I urged by telegram; still no answer. Finally, I tried to reach her by telephone at the Museum of Natural History In New York where she was a curator. I was told she was with the Papuas, or the Trobrianders, or the Samoans, and could not be reached. So, I had to edit her speech and invent a title. What struck me was her speaking about cybernetics in a cybernetical way. Thus I chose for her the title, “Cybernetics of Cybernetics.”

It appears to me today that the interest in the peculiar properties of concepts that apply to themselves, (or even need themselves to come into being) must then have been floating in the air. Francisco Varela, the Chilean neurophilosopher referred to them as “self-referential,” the Swedish logico-mathematician Lars Lofgren as “auto-logical.”

**Yveline:** If I were to ask you to give me the shortest description of the distinction between first-order cybernetics and second-order cybernetics, what would you say?

**Heinz:** I would say, first-order cybernetics is the cybernetics of observed systems, while second-order cybernetics is the cybernetics of observing systems.

**Yveline:** Very short indeed! Would you like to expand on this?

**Heinz:** Perhaps only briefly, because my “shortest description” is nothing else but a paraphrase of the description I made in my address, where I juxtaposed the two fundamentally different epistemological, even ethical, positions where one considers oneself: on the one hand, as an independent observer who watches the world go by; or on the other hand, as a participant actor in the circularity of human relations.

When taking the latter position, (the position I believe taken by systemic family therapists) one develops notions like “closure,” “self-organization,” “self-reference,” “self,” “auto-poiesis,” “autonomy,” “responsibility,” etc., etc. In other words, one arrives at the whole conceptual machinery of contemporary cybernetics, the cybernetics of observing systems, and thus one comes very close to the theme of your Congress: “Ethics, Ideologies, New Methods.”

**Yveline:** At the conclusion of your paper, *On Constructing a Reality*, which was published in Paul Watzlawick’s book *The Invented Reality*, you ask, “What are the consequences of all this in ethics and aesthetics?” You also wrote, “The ethical imperative: Act always so as to increase the number of choices.” And, “The aesthetical imperative: If you desire to see, learn how to act.” Can you add something to the connections between ethics, aesthetics and change; which from my point of view, are the three basic coordinates in family therapy?

**Heinz:** I like your three coordinates, because all three have a second-order flavor. And, of course, I am delighted that two of my imperatives correspond to two of your coordinates. However, I feel some uneasiness that your third coordinate “change” is not yet accompanied by an appropriate imperative. Let me remedy this situation at once by inventing an imperative for you; the therapeutic imperative: “If you want to be yourself, *change!*” Is this paradoxical? Of course! What else would you expect from change?

**Yveline:** You say with so much self assurance, “Paradoxical, of course!” How can you connect change with paradox?

**Heinz:** Easily! You remember paradox? It yields one meaning when apprehended one way, and one meaning when apprehended the other. What do you do when I say “I am a liar,” do you believe me? If you do, then I must have spoken the truth; but if I had spoken the truth, I must have lied, etc., etc. What is the problem here? Lying? No, the problem is “I,” the shortest self-referential loop. When speaking about oneself, using “I,” magic is performed. One creates oneself by creating oneself. “I” is the operator who is the result of the operation.

**Yveline:** This is all magic to me. Where does “change” come in?

**Heinz:** The paradoxical nature of change is much richer than the orthodox “paradox of the liar” which switches from “true” to “false,” and from “false” to “true,” and so on and so forth in dynamic stability. The unorthodox nature of change arises when you apprehend “change” any way you wish to apprehend it, and it will yield something else, otherwise it wouldn’t be “change.” This is, I believe, its therapeutic force.

**Yveline:** But you said, “If you want to be yourself; change!” How can you be yourself and change?

**Heinz:** I wanted to appeal to ancient wisdom. It is 2600 years old and comes from the *I Ching*. Under the 58th symbol “Fu,” or “The Turning Point,” it says, “The ultimate frame for change is the unchanging.”

**Yveline:** (Smiling) This conversation with you, Heinz, has been a joyful and exciting day of learning. It seems to have mirrored the theme of our conference; ethics and family therapy. It feels as though I’ve discovered a new freedom within a precise and rigorous framework. This framework, clearly defined by the fundamental guidelines of therapeutic practice, encourages communication with another, thereby creating a new space. Does this not broaden our possibilities by redrawing the line of the horizon? If rigor were combined with creativity, the ethics of choice could also be the ethics of change!

At least that is the very personal understanding which I have gained from our encounter. I now have an exquisite diffused feeling of a door which opens onto another door, which opens onto another door, which opens onto another door . . .