

Disorder/Order: Discovery or Invention?

Heinz von Foerster

1984

Ladies and Gentlemen: This is a great symposium. I enjoy every minute of it. However, I feel there is a blemish, and this is that Gregory Bateson is not with us. The reason why I, in particular, am sad he is not among us is not only because he would have enjoyed tremendously being here, and you would have enjoyed him very much as well, but because I need his help to put to rest one of the questions which has continually recurred during this conference. Here is the question: *Are the states of order and disorder states of affairs that have been discovered, or are these states of affairs that are invented?*

As I tend to say they are invented, I need all the help I can muster in order to defend this position, and so invoke the spirit of Gregory Bateson to stand on my side and to help me now for my defense. I will ask him to give us one of his very charming vignettes which pretend to be dialogues between a fictitious daughter and a fictitious father. (I do not think these fictions are too fictitious, indeed.) These dialogues he called *Metalogues*, and I will read you one now with a few comments on my side. This one is entitled *Metalogue: What is an Instinct?* It begins with daughter asking father, "Daddy, what is an instinct?" Now, if my daughter, or my son, had asked me, "Daddy, what is an instinct?" I, most likely, would have fallen into the trap of giving a learned, lexical definition. I, for instance, would have said: "An instinct, my dear, is the innate behavior of animals which is unlearned, has a certain complexity, etc.," or something like that. However, Bateson does not fall into that trap and, for an answer to "Daddy, what is an instinct?" he says: "An instinct, my dear, is an explanatory principle." That is not good enough for her; therefore, daughter immediately asks, "But what does it explain?" And he replies (now watch!) "Anything, almost anything at all, anything you want it to explain." Now, please realize, that something which explains "anything you want it to explain" of course explains nothing. But daughter immediately senses something, and she says, "Don't be silly. It doesn't explain gravity!" The father: "No, but that is because nobody wants instinct to explain gravity. If they did, it would explain it. We could simply say, 'The moon has an instinct whose strength varies inversely as the square of the distance. . .'" Daughter: "But this is nonsense, Daddy!"—"Yes, surely, but it was you who mentioned instinct, not I."—"All right, but what does then explain gravity?"—"Nothing, my dear. Because gravity is an explanatory principle." "Oh," says the daughter, "now, do you mean you cannot use one explanatory principle to explain another principle, never?" Father: "Hardly ever. That is what Newton meant when he said, *hypotheses non fingo*."—"And what does that mean, please?" asks daughter. (Now I would like to draw your attention to the fact that when the father gives his answer, everything that he says is put in the descriptive domain. It is always associated with saying or with pointing.) Again, daughter: "What does that mean, please?" Father: "Well, you know what hypotheses are. Any statement linking together two descriptive statements is a hypothesis. If you say there was a full moon on February 1, and another on March 1, and then you link these two descriptions together in any way, the statement which links them is a hypothesis."—"Yes, and I know what *non* means. But what is *fingo*?"—"Well, *fingo* is a late Latin word for 'make'. It forms a verbal noun, *fictio*, from which we get the word 'fiction'."—"Daddy, do you mean that Sir Isaac Newton thought that all hypotheses were just made up, like stories?" Father: "Yes, precisely that."—"But didn't he discover gravity? With the apple?"—"No, dear, he invented it!"

With this Batesonian dialogue I have, as it were, set the stage for what I am going to say. My original plan was to make some historical remarks in regard to the notion of disorder and order; however, during the development of

this conference, I realized I should indeed shift my emphasis. There were two points which persuaded me to do this: one, I realized that we have the tremendous pleasure of having Michel Serres here, who is one of the eminent historians and could of course say much better anything historical than I could ever invent; the second point is that I am not the last speaker, and since I feel that this conference has historical significance and what I will say today will be obliterated tomorrow, I am very happy that, in their wisdom, the organizers of this conference have put Michel Serres as the last speaker; moreover, I hope he will satisfy Edgar Morin's request that the observer include himself in the observation, for he would then also be a contributor to the history of this conference.

To what, then, am I to address myself when I am not addressing myself to history? I shall shift from the historical to the epistemological, because I have the feeling that many of the questions that have been raised during this conference have an epistemological root. Nevertheless, with your permission, I will make two points, where I will have osculations with historical events regarding the notions of disorder and order, and this is when our topic touches a certain branch of poetry, namely, thermodynamics. These points I shall discuss because I have seen that, again and again during this symposium, notions which developed from an interaction between people in the scientific fields let us say, the thermodynamicists and others, a *lingo*, a language, a notation, evolved, which is being used here, alas, in a somewhat loose fashion, and I would like to recall for you the occasion on which these notions arose. After I have made these brief contacts with history just to see the perspectives, I will then try to show that the notions of disorder, order, and organization are conceptually linked to a general notion of computation. This will give me a platform, first to talk in quantitative terms about order and complexity, hence of those processes by which order, or complexity, is increased or decreased; but secondly—and this is the essential justification for my tying these notions to computation—to show that these measures are fully dependent upon the chosen framework (which turns out to be the language) in which these computations are carried out. In other words, the amount of order, or of complexity, is unavoidably tied to the language in which we talk about these phenomena. That is, in changing language, different orders and complexities are created, and this is the main point I would like to make.

Since a free choice is given to us which language we may use, we have moved this point into a cognitive domain, and I will reflect upon two types of cognition which I already touched upon in my introductory statement; namely, the problem of whether the states that we call “disorder and order” are states of affairs that are discovered or invented. When I take the position of *invention*, it becomes clear that the one who invents is of course responsible for his or her invention. At the moment when the notion of responsibility arises, we have the notion of *ethics*. I will then develop the fundamental notion of an ethics that refutes ordering principles attempting to organize the other by the injunction, “Thou shalt,” and replace it by the organizational principle, that is, organizing oneself with the injunction “I shall.” With this note I have given you a brief outline of my talk. Now, ladies and gentlemen, I can begin with my presentation!

First, I would like you to come with me to the year 1850. This is approximately the time when the First Law of Thermodynamics was well established, one understood the principle of conservation of energy, and the, Second Law of Thermodynamics was just in the making. What was observed and what was intriguing people very much at that time was an interesting experiment. I ask you to look with me please at the following fascinating state of affairs. Consider two containers, or reservoirs, of the same size. One is hot, and the other one is cool. Now you take these containers, put them together, fuse them, so to speak, and watch what happens. Spontaneously, without our doing anything to them, the cold container will become warmer, and the warmer will become colder. Now, you may say, “O.K., so what?” But, ladies and gentlemen, if you say, “so what?” to anything, you will not see anything.

The engineers (and as Mr. Prigogine has so properly said, thermodynamics was an engineering science), who were working with steam engines, heat engines, etc., were wondering about the efficiency of these machines. They knew very well that if one has a hot and a cold container, one can put between these two vessels a heat engine that will do some work for us, drilling, pumping, pulling, and things like that. But they also knew that the smaller the temperature difference between these two containers is, the less the chance of getting a heat engine going; this means that the possibility of changing heat into work becomes less and less as the temperatures of the two containers become more and more alike.

When Clausius thought about that very carefully, he realized what is going on here: with the decrease in the difference between the two temperatures, the convertibility, the change, the turning of heat energy into work, becomes less and less possible. Therefore he wanted to give this possibility of being able to turn or to change heat into work a good and catchy name. At that time it was very popular to use Greek for neologisms. So he went to his dictionary and looked up the Greek for “change,” and “turn. He found the word *trope*. “Aha,” he said, “but I would like to talk about *not change*, because, you see, the longer these processes go on, the less heat can be turned into work.” Now unfortunately, either he had a lousy dictionary, or he could not speak Greek very well, or he had friends who did not understand what he was talking about. Instead of calling it *utropy*, because *ou* is the Greek word for *non*, as in “Utopia” (no place)—and *utropy* is what he should have called his new concept—for some reason he called it “entropy,” because he thought that *en* is the same as the Latin *in* and therefore means “no.” That is why we are stuck with the wrong terminology. And what is worse, nobody checked it! An incredible state of affairs! So, in proper *lingo*, when these two containers are put together, the *utropy* of the two increases, because the possibility for changing, for transforming the heat into work becomes less and less.

A couple of years later, two gentlemen, one in Scotland, one in Austria, one in Edinburgh, the other in Vienna, one by the name of Clerk Maxwell, and the other by the name of Ludwig Boltzmann, were intrigued by a fascinating hypothesis, a hypothesis which was so crazy that most of their colleagues in the academic community refused even to talk about that stuff. They were contemplating whether it would be possible to think of matter as not being indefinitely divisible, so that at a particular level of subdivision, one could not subdivide any further. That is, one would be left with small pieces of mass. “Mass” is *moles* in Latin, and for a small thing, one puts on the diminutive suffix, which is *-cula*, and we get the hypothetical, “molecules” that would not allow further division.

Contemplate whether this hypothesis makes any sense at all. To put you into the perspective of that time, 1871 or 1872, Boltzmann, who was teaching in Vienna, occupied one chair in physics. The other chair belonged to Ernst Mach, whose name, I believe, is familiar to you. Mach went into the Boltzmann lectures, sitting in the last row of the big physics auditorium, and when Boltzmann used the word “molecule” in his lectures, Mach screamed from the last row, “Show me one!” Of course, at that time one could not show one; they were purely hypothetical. Anyway, these two gentlemen, Maxwell and Boltzmann, addressed themselves to the problem of whether we can indeed interpret some of the fundamental laws of physics as if matter were composed of elementary particles, the molecules. They succeeded. They showed that three fundamental quantities in thermodynamics could be expressed in terms of molecular properties. The one is pressure. It is interpreted as a hailstorm of molecules flying against the walls of a container. The kinetic energy, or the speed of the molecules, would determine temperature. And then they came to the notion of entropy, or *utropy*, as I would say, and here a fascinating thing happened.

They could not explain *utropy* in purely molecular terms, and had to make an appeal to the cognitive functions of the observer. This is the first time when, in science, the observer enters into his descriptive system. What was necessary in order to handle the notion of *utropy*, was to talk about the distinguishability of states of affairs. I will give you an example. Take again the two boxes which can be distinguished by their different temperatures: one at a high temperature, the other at a low temperature. Put them together so that they are fused. Now the hotter will become colder, and the colder slowly warmer, and as time goes on their distinction will be lost: they become more and more “confused.” Better, the observer becomes “confused” because he will be unable to distinguish between the two containers, his confusion increasing with the increase of the *utropy*. Here you have one version of the Second Law of Thermodynamics: *utropy* increases with confusion. Or, as others may say: entropy increases with disorder.

Seeing the Fundamental Laws of Thermodynamics, which were originally formulated so as to account for a macroscopic phenomenology, to have—in turn—their foundation in a microscopic mechanics, stimulated questions about the potential and limits of these Fundamental Laws.

I can see Clerk Maxwell sitting there, dreaming up some mischief about how to defeat the Second Law of Thermodynamics: “Hmm, if I have two containers at equal temperature, what must go on between them so that, without external interference, the one gets hotter, while the other gets colder?” Or, if you wish, letting order (discriminability) emerge from disorder (indiscriminateness), i.e., reducing the entropy of the system. Maxwell, indeed, came up with a charming proposal by inventing a demon who would operate according to a well-defined rule. This demon

is to guard a small aperture in the wall separating the two containers and to watch the molecules that come flying toward this aperture. He opens the aperture to let a molecule pass whenever a fast one comes from the cool side or a slow one comes from the hot side. Otherwise he keeps the aperture closed. Obviously, by this maneuver he gets the cool container cooler (for it loses all its “hot” molecules) and the hot container hotter (for it loses all its “cool” molecules), thus apparently upsetting the Second Law of Thermodynamics. So, Maxwell invented his famous demon, whose name is, of course, “Maxwell’s Demon,” and for quite a while it was thought he would indeed have defeated the Second Law. (Later on, however, it was shown—but that is quite irrelevant to my story—that indeed, the Second Law of Thermodynamics is upheld, even with the demon working. Because in order for the demon to judge whether these molecules are fast or slow, he must of course have a flashlight to see these molecules; but a flashlight has a battery, and batteries run out, and there of course goes the hope of having defeated the Second Law of Thermodynamics!)

But there is another point that I would like to make regarding this demon, and that is that he is the incorporation par excellence not only of any principle that generates distinctions and order, but also of a general notion of computation. One of the most fundamental concepts of computation, I submit, was developed in the thirties by the English mathematician Alan Turing. He exemplified his notion with the aid of a fictitious machine, a conceptual device, the internal states of which are controlled by one, and are controlling the other one of the machine’s two external parts. The first one is a (theoretically infinite) long tape that is subdivided into equal-sized squares on which from a given alphabet (one may say “language”), erasable symbols can be written. The other part is a reading/writing head which scans the symbol on the square below it and, depending upon the machine’s internal state, will either change this symbol or else leave it unchanged. After this it will move to the next square, either to the left or to the right, and finally will change its internal state. When these operations are completed, a new cycle can begin, with the head now reading the symbol on the new square. In a famous publication, Turing proved that this machine can indeed compute all computable numbers or, as I would say in reference to our topic, all “conceivable arrangements.”¹

What I would like now to demonstrate is that this machine—whose name is, of course, the “Turing Machine”—and Maxwell’s demon are functional isomorphs or, to put it differently, that the machine’s computational competence and the demon’s ordering talents are equivalent. The purpose of my bringing up this equivalence is, as you may remember from my introductory comments, to associate with the notions of disorder, order, and complexity, measures that permit us to talk about different degrees of order, say: “More order here!” or “Less order there!”, and to watch the processes that are changing these degrees.

Let us now quickly go through the exercise of this demonstration by comparing the machine’s M and the demon’s D actions during the five steps of one complete cycle. Step (i): M reads symbol, D watches molecule; (ii): M compares symbol with internal state, D compares molecule’s speed with internal standard; (iii): M operates on symbol and tape, D on aperture, opening or closing it; (iv): M changes its internal states, D its internal standard; (v): M and D go back to (i). Q.E.D.

Knowing about this equivalence puts us in the position of transforming any ordering problem into a computational one. Consider, for instance, an arbitrary arrangement, A , and its representation on the tape of a Turing Machine by using a certain alphabet (language). What Turing showed is that there exists another tape expression, called the “description” of A , which, when used as the initial tape expression will allow the machine to compute from it the arrangement A . Let me now draw your attention to three measures (numbers). One is the length $L(A)$ (that is, the number of squares) of the tape that is taken up by the arrangement A ; the second is the length $L(D)$ of A ’s description (the initial tape expression); and the third figure is N , the number of cycles the machine has to go through to compute the arrangement A from its description D .

Now we can collect some fruits from our intellectual investment into the notions of machines, demons, etc. I will describe just four:

¹“On Computable Numbers with an Application to the Entscheidungsproblem,” in *Proceedings of the London Mathematical Society* 2, no. 42 (1936), 230–65.

(i) *Order*

If the initial tape expression, the description, is short, and what is to be computed, the arrangement, is very long ($L(D) < L(A)$), then it is quite obvious that the arrangement possesses lots of order: a few rules will generate A . Take A to be 0, 1, 2, 3, 4, 5, 6, 7, ..., 999,999, 1,000,000. A suitable description of this arrangement may be: *Each follower equals its precursor + 1*.

(ii) *Disorder*

If the length of the description approximates the length of the arrangement, it is clear that we do not understand this arrangement, for the description just parrots the arrangement. Take A to be:

8, 5, 4, 9, 1, 7, 6, 3, 2, 0.

I challenge the mathematicians present, or any puzzle wizard, to come up with a rule other than: *write 8, 5, 4, ... that generates this arrangement*.

(iii) *Complexity*

I propose to use N , the number of cycles for computing an arrangement, as a measure for the complexity of this arrangement. In other words, I suggest that we associate with the complexity of an arrangement the time it takes the machine to compute it. For instance, during this meeting a juxtaposition molecule/man was made with the suggestion—so I understood—to learn about the properties of human beings from the known properties of molecules. In computational jargon such computations are usually referred to as computations *ab ovo* or, as in our case *ab molecula*. From this point of view it may be not too difficult to see that N , the number of computational steps, will be so large (e.g., the age of the universe being too short to accommodate N) that N becomes “trans-computational.” That means, we can just forget about the whole thing, for we shall never see the end of it!

(iv) *Language*

The choicest of the four fruits I have left to be the last for you to taste, for it is the most crucial one in my narrative. It is the observation that all the three quantities mentioned before: the length of an arrangement, the length of its description, and the length of computing this arrangement, are drastically changed by changing from one alphabet a to another one, say, b . In other words, the degree of disorder or order that can be seen in an arrangement depends in a decisive way upon the choice of language (alphabet) that is used in these operations. Take as an example my telephone number in Pescadero. It is 879-0616. Shift to another alphabet, say, the binary alphabet. In that language my number is 100001100010001001011000. Should you have difficulties remembering that number, shift back to the former language!

Take as another example the random number sequence 8, 5, 4, etc., I spoke of earlier (point *ii*). I suggest shifting from an alphabet that uses Arabic numerals to one that spells out each numeral in English: 8—eight, 5—five, 4—four, etc., and it becomes clear that under this alphabet the former “random sequence” is well determined, hence has a very short description: it is “alphabetical” (eight, five, four, nine, one, etc.).

Although I could go on with a wealth of examples that would drive home again and again the main points of my argument, in the hope that the foregoing examples suffice I will summarize these points in two propositions. Number one: A computational metaphor allows us to associate the degree of order of an arrangement with the shortness of its description. Number two: The length of descriptions is language-dependent. From these two propositions, a third one, my essential clincher, follows: Since language is not something we discover—it is our choice, and it is we who invent it disorder and order are our inventions!²

With this sequence I have come full circle to my introductory claim that I shall once and for all put to rest the question of whether disorder and order are discoveries or our inventions. My answer, I think, is clear.

Let me draw from this constructivist position a few epistemological consequences that are inaccessible to would-be discoverers.

²Except for the Greeks, who believed that it was the Gods who invented language and that we humans are doomed to discover it.

One of these is that properties that are believed to reside in things turn out to be those of the observer. Take, for instance, the semantic sisters of Disorder: Noise, Unpredictability, Chance; or those of Order: Law, Predictability, Necessity. The last of these two triads, Chance and Necessity, have been associated until even recently with Nature's working. From a constructivist point of view, Necessity arises from the ability to make infallible deductions, while Chance arises from the inability to make infallible inductions. That is, Necessity and Chance reflect some of *our* abilities and inabilities, and not those of Nature.

More of that shortly. For the moment, however, let me entertain the question of whether there exists a biological backup for these notions. The answer is yes, and indeed, I am very happy that we have just those people around who were producing this very backup that allows me to speak about an organism as an autonomous entity. The original version came from three Chilean neuro-philosophers, who invented the idea of autopoiesis. One of them is sitting here, Francisco Varela; another one is Umberto Maturana, and the third one is Ricardo Uribe, who is now at the University of Illinois. They wrote the first paper in English on the notion of autopoiesis, and in my computer language I would say that autopoiesis is that organization which computes its own organization. I hope that Francisco will not let me down tomorrow and will address himself to the notion of autopoiesis. Autopoiesis is a notion that requires systemic closure. That means organizational, but not necessarily thermodynamic, closure. Autopoietic systems are thermodynamically open, but organizationally closed.

Without going into details I would like to mention that the concept of closure has recently become very popular in mathematics by calling upon a highly developed branch of it, namely, Recursive Function Theory. One of its concerns is with operations that iteratively operate on their outcomes, that is, they are operationally closed. Some of these results are directly associated with notions of self-organization, stable, unstable, multiple and dynamic equilibria, as well as other concepts that would fit into the topic of our symposium.

However, traditionally there have always been logical problems associated with the concept of closure, hence the reluctance until recently to take on some of its problematic aspects. Consider, for example, the relation of an observer to the system he observes. Under closure, he would be included in the system of his observation. But this would be anathema in a science where the rule is "objectivity." Objectivity demands that the properties of the observer shall not enter the descriptions of his observations. This proscription becomes manifest when you submit to any scientific journal an article containing a phrase like "I observed that—." The editor will return it with the correction "It can be observed that..." I claim that this shift from "I", to "it" is a strategy to avoid responsibility: "it" cannot be responsible; moreover, "it" cannot observe!

The aversion to closure, in the sense of the observer being part of the system he observes, may go deeper. It may derive from an orthodox apprehension that self-reference will invite paradox, and inviting paradox is like making the goat the gardener. How would you take it if I were to make the following self-referential utterance: "I am a liar." Do I speak the truth? Then I lie. But when I lie, I speak the truth. Apparently, such logical mischief has no place in a science that hopes to build on a solid foundation where statements are supposedly either true or else false.

However, let me say that the problems of the logic of self-reference have been handled very elegantly by a calculus of self-reference, whose author is sitting on my left (Varela). I hope he will not let me down and will give me a bit of self-reference when he speaks tomorrow!

Social theory needs agents that account for the cohesiveness of social structure. Traditionally the agents are seen in sets of proscriptions issued with some dictatorial flavor, usually of the form "Thou shalt not..." It is clear that everything I said tonight not only contradicts, but even refutes, such views. The three columns, autonomy, responsibility, choice, on which my position rests, are pointing in the opposite direction.

What would be my counter-proposal? Let me conclude my presentation with a proposition that may well serve as a Constructivist Ethical Imperative: "I shall act always so as to increase the total number of choices."

Discussion

PAUL WATZLAWICK [STANFORD]. Heinz, would you say that, in addition to what you call the “ethical imperative,” there is still a further conclusion to be drawn, and that is that if you realize that you are the constructor of your own reality, you are then also free, and so the question of freedom enters, so there is a deontic quality to what you were saying?

VON FOERSTER. My response is: Yes, precisely.

KARL H. PRIBRAM [STANFORD MEDICAL SCHOOL]. Heinz, I agree with everything you said, and with what Francisco says, but I have a problem. And that problem is, given the kind of framework you have just “invented” for us, and which I like very much, why is it that when I go into the laboratory, something happens that surprises me? When I know how things are supposed to go, and they don’t.

VON FOERSTER. You are a very inventive character—you even invent your surprises. For instance, when I was talking about the two containers that are brought together and said that a most surprising thing is taking place, namely, that the hot one is getting cooler, and the cool one getting hotter, I felt that apparently this was seen as a joke—of course, everybody knows that, so what? But my hope was that you would try to see this phenomenon as if for the first time, as something new, something fascinating. Let me illustrate this point. I don’t know whether you remember Castaneda and his teacher, Don Juan. Castaneda wants to learn about things that go on in the immense expanses of the Mexican chaparral. Don Juan says, “You see this . . . ?” and Castaneda says “What? I don’t see anything.” Next time, Don Juan says, “Look here!” Castaneda looks, and says, “I don’t see a thing.” Don Juan gets desperate, because he wants really to teach him how to see. Finally, Don Juan has a solution. “I see now what your problem is. You can only see things that you can explain. Forget about explanations, and you will see.” You were surprised because you abandoned your preoccupation with explanations. Therefore, you could see. I hope you will continue to be surprised.