*It's not easy to alter engrained behavior even when that behavior is clearly
counterproductive – but, fortunately for the health
of our organizations, it is possible.*

The Executive Mind and

Double-Loop Learning

Chris Argyris

Over the past decade, I have been studying the ways executives reason while they are solving difficult human and technical problems. The executive mind seems to work in bewildering ways, a few of which I will discuss in this article. For one thing, I have identified a pattern of three nested paradoxes embedded in executive reasoning.

First, the reasoning executives use to manage people and technical issues leads simultaneously to productive *and* to counterproductive consequences.

Second, they are unaware of this feature because they are disconnected from their own reasoning processes while making tough decisions.

Third, they are disconnected from their reasoning processes because of the skills they have mastered to solve tough problems. The skills that lead to success will also lead to failure.

How can the same reasoning necessarily lead to productive and counterproductive consequences? How can people act and at the same time be disconnected from their reasoning processes? And why is it necessary

Chris Argyris *is James Bryant Conant Professor of education and organizational behavior at Harvard University. His academic degrees are: A.B., Clark University; M.A., Kansas University; Ph.D., Cornell University; and honorary doctorates from McGill University, the University of Leuven, Belgium, and the Stockholm School of Economics. From 1951 to 1971, Argyris was Beach Professor of administrative sciences at Yale University; he served as chairperson of the department during the latter part of the period. His early research focuses on the unintended consequences for individuals of formal and organizational structures, executive leadership, and control and management-information systems — and on how individuals adapted to change those consequences* (Integrating the Individual and the Organization, *John Wiley & Sons, 1964*). *He then turned his attention to ways of changing organizations* (Organization and Innovation, *Irwin-Dorsey, 1965*). *This led him to focus on the role of the social scientist as researcher and interventionist* (Inner Contradictions of Rigorous Research, *Academic Press, 1980*). *During the past decade, he has been developing, with Donald Schön, a theory of learning in which human reasoning, not just behavior, becomes the basis for diagnosis and action* (Organizational Learning, *Addison-Wesley, 1978*). *Argyris's latest book is* (Reasoning, Learning, and Action: Individual and Organizational *(Jossey-Bass, 1982)*.

for them to be disconnected from their reasoning processes in order to solve difficult problems? What impact do these features have on executive problem solving and on the organization?

These nested paradoxes indicate that we are dealing with some deeply em-

bedded features of the human mind. And it is the executive mind that concerns us because executives are most often held responsible for dealing with the difficult issues in organizations and in society at large.

The Executive Mind Defined

By "executive mind," I mean the way executives create premises, make inferences, and arrive at conclusions. Surprisingly, executives (or anyone else, for that matter) are usually unaware of their reasoning processes. There are two reasons for this. First, they have great reasoning skill — the activity is second nature to them, and they are rarely aware of it while they are doing it. Indeed, as is true of most skilled behavior, they rarely focus on it unless they make an error. Second, when they do make errors, other people — especially subordinates — may feel it is safest to play down the error, or may ease in the correct information so subtly that the executive will probably not even realize that he did make an error.

These actions at the upper levels are especially detrimental to the organization's capacity to detect and correct errors, to innovate, to take risks, and to know when it is unable to detect and correct error. Such consequences can lead to difficulties in getting the everyday job done correctly. But worse, lack of attention to the underlying policy issues can lead to the organization's losing control of its destiny.

In a previous article I gave an example in which managers at all levels "rounded out" sentences in reports in ways designed to (1) avoid upsetting those at the top too much and, at the same time, (2) permit the subordinates involved to "cover" themselves. The result was a multimillion dollar error that led to the closing of major facilities. The error was known, and the consequences predictable, by managers at the

Figure 1

WHAT Y SAID TO X

1. X, your performance is not up to standard (and moreover. . .)

2. You seem to be carrying a chip on your shoulder.

3. It appears to me that this has affected your performance in a number of ways. I have heard words like *lethargy, uncommitted,* and *disinterested* used by others in describing your recent performance.

4. Our senior professionals cannot have those characteristics.

5. Let's discuss your feelings about your performance.

6. X, I know you want to talk about the injustices that you believe have been perpetrated on you in the past. The problem is that I am not discussing something that happened several years ago. Nothing constructive will come from it. It's behind us.

7. I want to talk about you today and about your future in our system.

---

lower levels several years before the crisis exploded into the open.

Our research indicates that when executives deal with difficult, threatening, underlying issues, they use reasoning processes that, at best, simultaneously lead to immediate success and long-range problems. Often the problems go unsolved, compounding the long-range difficulties. Much of this occurs without executives' realizing it. Or, if they do realize it, many believe that no other outcome is possible. They are correct if they are willing to accept the world as it is, without seeking alternatives. According to the Pentagon Papers, this is what happened at upper levels in the Defense Department as well as in the State Department.

To illustrate how we arrived at these conclusions, I want to present data from a case that concerns one of the most difficult problems executives face—namely,

to help fellow executives realize that their performance is deteriorating when they believe otherwise. Later I will show how the results from this case apply to other common but difficult leadership problems. (The reader may wish to try his hand at solving this case and compare his or her response with those of our sample.)

CORRECTING POOR PERFORMANCE: A DIAGNOSTIC CASE

Y, a senior executive, must tell X, an older officer, that his performance during the past five years has fallen below standard. Y knows that the difficulty of his task is compounded by the fact that X believes his performance has topped off because of the way the firm has dealt with him.

We give the executives in a seminar a transcript of several key sentences that Y used in talking with X (Figure 1) — sentences that represent the range of meanings that Y communicated to X during their session.

We then ask the executives to answer three questions:

1. How effective do you believe Y was in dealing with X?

2. What advice would you give Y?

3. Assume that Y asked you for your evaluation of his effectiveness in dealing with X. Write your response, using the following format, for two or three double-spaced pages.

Your thoughts and feelings	What you and Y said
(Give in this column any thoughts and feelings you had during the session but which you did not communicate.)	I: (Write what you would say)
	Y: (Write what you expect Y's response would be.)
	I: (Write your response to this.)
	Y: (Write Y's response.)
	and so forth

The responses presented below 7

were made by the 15 top senior officers, including the CEO, of a five-billion-dollar corporation. If your answers do not vary significantly from those in our sample, then the consequences that followed for the executives will more than likely occur for you. (If your answers are different, please mail them to me. It will help us better understand the gaps in our research.)

The Executives' Evaluation of Y's Effectiveness

Figure II is a collage of the executives' answers to Question 1. Briefly, the results are as follows:

• The executives evaluated Y as being ineffective in his dealings with X.

By the way, the reader may be interested to know that line executives were as compassionate about X as were governmental executives and organization development professionals. Indeed, the line executives were slightly more concerned than the other two groups that Y was too power-oriented, an "uncaring executioner of company policies."

• The executives organized their responses by inventing what might be called a miniature causal theory of human behavior: *If* Y (or anyone else) communicates meanings of the kind that Y communicated to X, *then* the recipient will feel defensive and learning will be blunted.

If you agree with this explanation, then you are using an explanation that, strictly speaking, doesn't come from the data. For example, why should telling X that his performance is poor and unacceptable make him defensive? "That is obvious," you may respond, "because such statements are probably experienced as punishing and unjust, an attack on the person's competence. Such acts are threatening." To arrive at this

conclusion — which is probably correct — you must hold a tacit theory of threat.

• Therefore, embedded in the first causal miniature theory is another one, a theory about what makes people defensive.

• All the explanations that the executives produced have a taken-for-granted quality because they have been absorbed and learned in the experience of everyday life. In my jargon, people are socialized to believe this theory of defense because they have been taught from an early age that such an act — telling someone his or her performance is poor — can be counterproductive.

Some Puzzles and Surprises

First, two puzzles:

1. If everyone has learned these causal explanations early in life, why have they not also eventually learned *not* to create them? Why, for example, did Y behave the way he did toward X?

2. If Y had asked these executives what they thought of the way he dealt with X, and if the line executives told Y what is in Figure 2, they would be using the same causal theory with Y that they were criticizing Y for using with X. For example, to tell Y that he is insensitive and blunt is to be blunt and insensitive.

When the executives were confronted with the first puzzle, they were quick to reply that either Y was not very competent, or that maybe he was even a bit flustered and tried to cover this up by being directive.

Notice the reasoning. The executives dealt with the first puzzle (of why Y behaved thusly) by saying that Y lacked some skill, or was upset. That places the responsibility on Y. It also means that the executives do not have to question the validity of their diagnoses.

8

How accurate are their diagnoses? Let us ask Y.

When the executives said that Y:	Y could have responded that he:
1. Blamed X completely.	1. Blamed X justly, or only partially.
2. Resisted hearing X's views.	2. Only resisted getting into a past history that was a can of worms.
3. Discounted X's feelings.	3. Understood X's feelings for X's sake, but, he did not want to get mired in them; that he was trying to think positively and look toward the future.

When the executives were confronted with the second puzzle (that they were using the same approach as Y), their initial response was surprise and disbelief. A few tried to prove that this was not an inconsistency on their part, but fellow executives disagreed. When the defensive reactions wore thin, there was a momentary silence. Then someone said:

> Yes, you are right; there is the inconsistency. But what you fail to realize is that none of us would say to Y what we have written down.

> No, added another with a smile, we're too smart to say what we think.

Note the difference between the executives' reactions to the first and second puzzles. Their response to the first puzzle was to place the blame on Y. In the second, unable to explain away their own inconsistency, they decided that they would not say what they thought. They accepted being surprised about, and unaware of, their impact on Y as natural. But note that they never suggested that Y might have been unaware of his impact on X.

Once the executives had made a diagnosis, they assumed that it was true and countered any questioning of it in ways designed to keep their diagnosis intact.

Figure 2

ANALYSIS OF RESPONSES TO X AND Y CASE
(by line executives)

Y's Action Strategies

Y's comments have a strong power tone; they smell of conspiratorial knowledge.

Y gives no sign of interest or compassion.

Y set X up to give only the answers Y wants to hear.

Y comes across as a blunt, uncaring executioner of the firm's policy decision with regard to X.

Y makes it abundantly clear that he does not want to be bothered with X.

Y is insensitive to X's feelings.

Y waits too long to listen to X.

Y does not give X a chance to respond. He pays lip service to hearing X's side.

Y is too blunt, direct, one-sided.

Y cuts off X.

Y communicates the seriousness of the situation from the company's point of view.

At no time does he appear to communicate that he and the firm genuinely want X to have a second chance.

Impact on X

Makes X feel defensive, rejected.

Makes X defend his past performance aggressively.

X will feel prejudged, as though the "deck" were "loaded."

X is not likely to relax and learn.

X is left with no room for constructive exploration.

X feels totally demoralized and inferior.

X is placed in a no-win situation.

Impact on Learning

Inhibits learning on part of both individuals.

X will probably look for another job.

THE EXECUTIVES' ADVICE TO Y

Recall that the executives tried to evade the second puzzle (that they were using the same causal theory used by Y) by asserting that they would not tell Y their diagnosis.

9

We collected two kinds of data to test their assertion. First, the executives were asked to write scenarios of a conversation they might have with Y about Y's handling of X. They were at liberty to mold Y's reactions as they wished.

The executives used three basic strategies in their scenarios. The first strategy (used by five executives) was direct. In effect, they told Y their diagnosis. But many who used a direct approach communicated with Y in such a way that they neglected important features of their own advice. For example, most of them were "up front" with Y but were unable to create an "open process," "a receptive mood," a process for the "growth and development" of Y.

Here are examples of how three of these executives began their scenarios, with hypothesized reactions that Y might have had but did not disclose.

Examples of the Direct Approach

Executive said	According to the executives, Y could have thought
Executive 1: To tell the truth, I don't think you accomplished what you wanted to.	1. He is not only telling me I failed, but that I was blind to that fact.
Executive 2: Some of your comments were bound to hit X pretty hard and force a defensive reaction.	2. I had to be straight with him. Speaking of hitting pretty hard, you're not doing badly yourself!
Executive 3: Well, you started off pretty hard on his performance, his defensive reaction. The poor bastard almost had to defend his record to set you straight.	3. Somebody had to be straight with him. We had been pulling our punches for five years. He didn't have much to defend, and he knew it. I feel I have to set you straight about me!

A more lengthy example from one scenario indicates the flow of conversation and the possible buildup of Y's reactions.

Executive said	According to writer, Y could have thought
4. Nobody believes the	4. If this is true, I really was

judgment about their own poor performance is accurate unless they can balance it against a number of successes.

Did you get a defensive reaction from X?

Well, my only point is that you probably were guaranteed to get some defensiveness.

Don't be afraid to be honest.

doomed to failure because this guy has no successes.

Are you also thinking the same about me? If so, how about a few successes?

Of course I did, just as you predicted from what you just said.

Just like you are guaranteed to do with me. Let's begin to end this diplomatically. ("Thanks for the advice.")

I'm not afraid to be honest except with people like you.

To summarize the findings from the scenarios:

The executives	The executives advised Y
1. Were in unilateral control.	1. Not to unilaterally control X.
2. Evaluated Y unilaterally and negatively.	2. Not to evaluate X.
3. Failed to hear Y's pent-up feelings.	3. To listen to X's pent-up feelings.
4. Failed to solicit Y's participation.	4. To solicit X's participation.

So a third puzzle was that the executives created the very conditions they advised Y not to create with X, *and* they appeared to be unaware of doing so.

The Easing-in Approach

Another approach we might call an easing-in approach (utilized by ten executives) or a combination of a *beginning* easing-in approach and a *final* direct approach. The easing-in approach basically asked Y questions in order to get him to see his errors.

Executive said	Y could have thought
1. It's hard to say from just reading your notes. I guess I would like to hear a little about how X reacted.	1. What is he driving at? Why does he want to focus on X's reaction?

2. What do you think he was thinking?	2. How do I know what X was thinking? I told him that X didn't say much. What is he driving at?
3. Do you suppose there might have been a way to let him know you mean it?	3. I did let him know that I meant it. If he didn't believe me, that's more his problem than mine. I think this guy has an agenda.
4. You mean his attitude wasn't any better in the meeting?	4. Couldn't you tell that he was pretty unresponsive? Again, what is he driving at? Does he believe that X's attitude was better?
5. Perhaps if you could persuade him to open up about it, he might get it off his chest.	5. Maybe this is what he is driving at. If so, he is wrong. The last thing I want to do is open up past wounds . . . oh, these bleeding hearts!

Although the executives asked Y questions, if our interpretations about the impact on Y are valid, they were acting in ways that placed them in control of Y.

To summarize:

The executives	The executives advised Y
1. Were in unilateral control.	1. Not to unilaterally control X.
2. Evaluated Y as ineffective.	2. Not to prejudge X as being ineffective.
3. Made attributions about Y's motives.	3. Not to make attributions about X's motives.

Again the executives created the very conditions that they advised Y not to create with X, *and* they appeared unaware of having done so.

We conclude that when the executives tried to communicate information that they believed was threatening, they unknowingly created conditions of miscommunication, misunderstanding, and inconsistency. None of them tested their views openly. This led to self-fulfilling prophecies, because every comment made by Y in his defense they saw as validating their diagnosis. This made Y even more defensive, which "proved" to the executives that their diagnosis was correct. Since the executives never publicly tested their views, they did not know that they had created a self-fulfilling prophecy. So we not only have a self-fulfilling prophecy but one that is self-sealing as well!

We have obtained these results with 27 groups, encompassing slightly under 1,000 subjects of varying ages, positions, roles, cultures, and of both sexes. In all cases, we taped the discussions, in which many different views were expressed, as well as feelings of bewilderment and frustration. Analysis of tape recordings of these discussions also illustrate that when the participants disagreed with each other, they did so by using the same counterproductive reasoning and strategies just described.

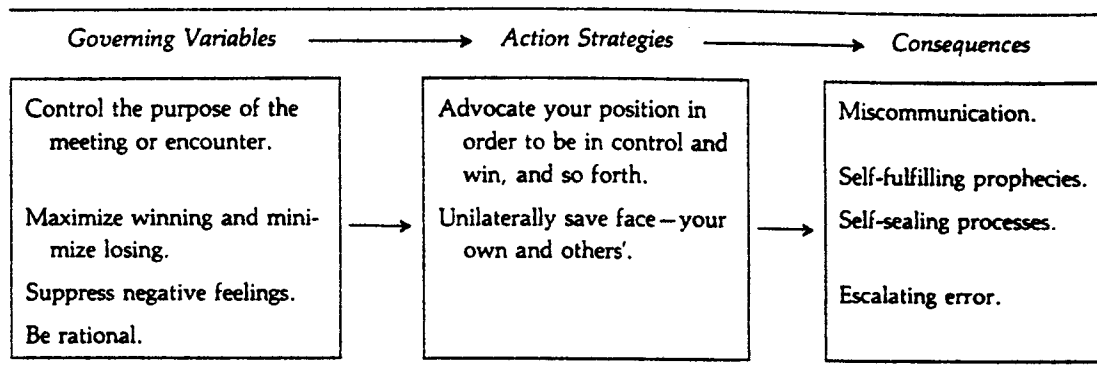## How Do We Explain These Results? Model I Behavior

In a previous article, I suggested an explanation for such results. Briefly, people acquire through socialization two kinds of skills and values for dealing with other people. The first are the values and skills that they espouse, the ones of which they are conscious and aware. I call these *espoused* theories of action.

Often when people are dealing with difficult and threatening problems, their behavior is inconsistent with their espoused theories. "Do as I say, not as I do" illustrates the point and at the same time proves that the point is not new.

What is new is the idea that all behavior is designed in accordance with a theory that we actually use. Moreover, we are rarely aware of this type of theory of action because it is ingrained in us from early childhood. I call it the *theory-in-use*. We use it without thinking about it. When we do think 11

<figure>

Figure 3

Model I Theory-in-Use

Governing Variables ———→ Action Strategies ———→ Consequences

| Control the purpose of the meeting or encounter.

Maximize winning and minimize losing.

Suppress negative feelings.
Be rational. | Advocate your position in order to be in control and win, and so forth.

Unilaterally save face — your own and others'. | Miscommunication.

Self-fulfilling prophecies.
Self-sealing processes.

Escalating error. |
</figure>

about it, we see that the results are often at odds with what we espouse. For example, the executives' espoused theories advocated dealing with Y in such a way that he did not become defensive. Many followed an easing-in approach; the theory-in-use for easing in is to ask Y questions which, if he answers correctly, will enable him to discover what we are hiding. As we have seen, this is a strategy that can be counterproductive and, in fact, the other person may imitate this approach by not saying what he or she really feels.

Although each person varied in what he said when using an easing-in approach, there was almost no variance in people's theory-in-use. We have observed the same theory-in-use among rich and poor, white and black, male and female, young and old, powerful and powerless, and in several different cultures.

We have created a model of the theory and call it *Model I* (Figure 3). It is composed of governing values, or variables, and action strategies and consequences.

Now we can return to the question of why we get such consistent results in the X and Y case.

Evaluations and attributions the executives made about Y's motives were not

self-evident. They required several layers of inference. For example, Y's statement to X, "Your performance is below standard," and so forth, is on the first rung of a ladder of inference. The second rung is the culturally understood meaning of such a sentence: "X, your performance is unacceptable." The third rung up consists of the meanings that the executives imposed. Here is where their theory-in-use came into play. They explained Y's actions by describing him as insensitive, blunt, and not listening.

Why do people use concepts at such a high level of inference (Figure 4)? The environment in which we operate is significantly more complex than what the human mind can process at a given moment. In order for the human mind to deal with reality, we must abstract from the buzzing confusion of everyday life (rung 1) by using more abstract concepts.

There are two key features of these concepts. First, they can be used to cover a lot of different meanings. For example, "blunt" and "insensitive" can apply to many different sentences. A second feature is that they are usually learned at an early age. Hence most of us learn to use the same concepts in similar ways. Soon we take them for granted. They become obvious and con-

12

Figure 4

LADDER OF INFERENCE

4	Meanings imposed by the researcher.
3	Meanings imposed by our theories-in-use.
2	Culturally understood meanings.
1	Relatively directly observable data, such as conversations.

crete, not abstract and questionable. Recall that there was a high degree of consistency and agreement in the way the executives diagnosed Y's actions toward X. It was "obvious" to most of the executives that Y had behaved counterproductively.

But the features of the human mind that make it efficient may also lead to counterproductive consequences. Why? Because, as we have seen, there can be differences between the executives' views and Y's. But so what? Cannot human beings differ in their views? Of course they can. But if they wish to reduce the number of misunderstandings when they do differ, they should test the validity of their reasoning.

Why do people choose not to test the validity of their reasoning? First, because they believe their reasoning is obvious and correct. Second, because they hold a theory of human defensiveness that tells them that the recipient will probably not listen, or will probably become defensive if he does. And it makes little sense for them to test their views with someone who, they believe, will be defensive.

Third, in a Model I world, testing makes the testers vulnerable. They could discover that they are wrong! They could lose unilateral control and they could generate negative feelings, consequences that would violate their Model I governing values.

The same counterproductive reasoning and consequences have been shown to occur when executives deal with technical and organizational problems. The only requirement is that the problems are threatening to some of the key players involved. Once the players diagnose the problem as threatening, they use the same reasoning described in the X-Y case.

Other Examples of the Consequences of Model I Reasoning

People programmed to be in unilateral control and to maximize winning tend to have difficulties in dealing with paradoxes. Paradoxes contain contradictions, and holding contradictory views makes the actor vulnerable to criticisms of being vague or self-contradictory. It is not surprising to learn from Louis B. Barnes's insightful descriptions that the way executives deal with paradoxes may be counterproductive.

Barnes's observations	Theory of action explanation
Often we fail to go beyond our initial reactions in order to look at deeper levels of the issue.	Do not run the risk of losing control and making yourself vulnerable — that is, losing.
Issues fall in opposing camps; hard data and facts are better than soft ideas and speculation.	Create win/lose dynamics. Seek hard data to win, to prevent losing. Abhor speculation lest you become vulnerable.

Turning to organizational consequences, recall that people who use Model I tend, as we have seen, to create misunderstanding, self-fulfilling prophecies, self-sealing processes, and escalating error. This drives people to follow Model I behavior even more closely in an attempt to maintain control, to win. It also creates competitive win/lose group and intergroup dynamics with many protective games that are undiscussable. And that undiscussability is itself undiscussable. The Model I world is, as the executives reported to Barnes, "an unsafe place where nice guys finish last."

An excellent illustration of these 13

consequences at the group, intergroup, and interagency level can be found in a recently published story of the interaction among Secretary of Health, Education, and Welfare Joseph A. Califano, Jr., Secretary of Labor Ray Marshall, and President Carter and his aides regarding welfare reform. The players acted toward each other in Model I ways. Califano kept secret for a long time his doubts of genuine welfare reform at zero cost increase. The President sensed his doubts but apparantly never explored them directly with Califano. When Califano said that he was working hard on a plan that Carter might possibly be able to call his own, the President asked Califano for a plan that he would gladly call "the Califano Plan." Califano's policy analysts were frustrated by his actions. They never said so, but they built up strategies to protect themselves. These strategies got them in trouble with Califano and a competing group of analysts in Secretary Marshall's office. This led to interagency warfare, a state in which positions harden and everybody looks out for Number One.

Robert B. Reich's recent article suggests that an entire industry is rising as a result of the self-sealing prophecies, self-sealing processes, and escalating error between private and governmental sectors. The new industry is composed of experts who deal with difficult relationships of private business and government by:

• Seeking to achieve clear controversies in which the client's position can be sharply differentiated from that of its regulatory opponent.

• Exaggerating the danger of the opponent's activities.

• Prolonging and intensifying conflict.

• Keeping business executives and regulatory officials apart.

Remember that regulatory agencies are also administered by Model I reasoning processes. This means that the regulators will probably deal with difficult, threatening

14

issues that are undiscussable by translating them into discussable, nonthreatening issues. For example, I have found that if regulators do not trust builders, instead of dealing with that issue they create piles of regulation in an effort to prevent cheating by dishonest builders. But these regulations may drive out the honest builders while stimulating the dishonest ones to new heights of creative dishonesty.

To summarize, holding a Model I theory-in-use makes it highly likely that the reasoning used for any difficult threatening issues, whether technical or personal, whether at the individual, goal, intergroup, organizational, or interorganizational level, will have counterproductive features that lead to self-fulfilling, self-sealing, error-escalating processes.

The reasoning is the same because it is individuals who deal with the substantive or human problems; who act as agents for groups, intergroups, organizations, and interorganizations. Programmed with Model I, they seek to win and not lose, to be in unilateral control, to suppress negative feelings. Thus whether the issue at hand is helping Y realize his error or dealing with a group that the executive believes is recommending the wrong investment strategy, the executives in either case will try to communicate their views in such a way that they cannot be held responsible for upsetting Y or the group members. The executive strives to evade responsibility for the defensiveness of others so that they cannot attack him and ignore the validity of his views. But, as we have seen, the strategies executives used to minimize the potential accusation of making others defensive actually does make others defensive, but in a way that makes it difficult or unlikely for them to say so. We then have the appearance of agreement.

Not only do we find these same consequences at all levels of the organization, we also find that managerial policies

and practices are designed to take these consequences into account. For example:

1. To reduce the probability that individuals will be able to blame the superior's evaluation of their performance, have them list and sign ahead of time a set of specific goals that are objectively measurable if possible.

2. To reduce the probability of group think, have several competing groups deal with the same problem.

3. To reduce miscommunication about difficult issues, have people send each other detailed position statements, backed up with hard numbers.

4. To reduce possible misunderstanding between regulators and those who are regulated, have the regulators define in detailed, unambiguous terms the standards of acceptable performance.

The first policy requires a technocracy of MBO experts and trainers. The second duplicates efforts. The third requires staff groups that may "pencil and paper an issue to death." The fourth requires mountains of specifications, including specifications on how to understand the specifications.

What Can Be Done

What can interested readers do to begin to learn more about the effectiveness of their reasoning processes and their action?

First, may I remind the reader of the two most fundamental findings. They were that the executives' responses were highly automatic and skillful, and that the executives were programmed to be unaware of their faulty reasoning as well as of the counterproductive impact. To change highly skilled action is not easy; to do so hampered by programmed unawareness is difficult indeed. And when the basis for the programmed unawareness is what we have been taught since early childhood, the task becomes formidable.

However, it is by no means an impossible task. The reader is not starting from scratch. We all know how to reason to separate effective from ineffective results, and how to design experiments for learning. Our research suggests that it will take about as long to learn the new reasoning and action skills as it takes to learn to play a good game of tennis or golf. And, in my opinion, that is how it should be. We are talking about changing our reasoning processes, one of our most fundamental human features. The learning should be difficult in order to rule out the gimmick hunters and quick-fix seekers.

Learning new reasoning processes, like learning to play tennis or golf, requires plenty of practice. Herein lies an advantage because most of us are constantly in situations in which we must use reasoning processes. There is plenty of opportunity to practice in everyday life.

The nature of the practice will be different at different stages of learning. During the early stages, you should be able to make errors without a high cost to you or to the organization. Also, you should have plenty of time to get feedback and to redesign your actions. These two features, when combined, suggest that the best kind of learning environment during the early stages is one that allows for slowdown of the action and decomposition of the problem. This in turn means creating a learning environment separate from everyday pressures.

For example, select a double-loop issue that is important to you, one that requires you to deal with others in order to solve it. Using the X-Y case format, write in one paragraph how you define the issue. In a second paragraph, write how you tried to solve it (or might try to solve it if it is a future problem). Next, write an actual scenario of several pages describing the conversation as you can best recall it (or if it is a future problem, what you would expect the conversation to be). Include thoughts and 15

feelings that you might not communicate, for whatever reason.

By the way, do not worry about how accurately you recall the incident, or how well you plan the future dialogue. If our theory is correct, you cannot write down anything except what is consistent with your theory-in-use.

Now put the case away for at least a week. When you reread it, analyze it as if you were trying to help a friend. Here are some of the questions that you can ask yourself about the dialogue.

• Do the sentences indicate advocating a position in order to be in control and to win and not lose? Or is the advocacy of the position combined with encouraging the other person to inquire? Is there an easing-in or forthright strategy? How aware is the writer of the possible interpretations by the receiver?

• Are the evaluations or attributions made with or without illustrations? Are they tested publicly or do they go untested?

• What kind of information is on the left-hand side of the paper (thoughts and feelings)? Does it contain information that would better enable the other person to understand your intentions? If so, what prevented you from communicating this information.?

• If the feelings and thoughts in the left-hand column would predictably upset others, what change would be necessary so that they could be effectively communicated?

More important, why does the writer think and feel about other people in ways that are not directly communicable? Sure, it may be that they are S.O.B.s. But it may also be that the writer is unknowingly creating self-fulfilling and self-sealing processes.

The next step would be to try to redesign some part of the dialogue, especially the sentences with which you find it difficult to deal.

16 • Read the sentence(s) several times and

write down the (culturally acceptable) meaning that you infer (rung 2 on the ladder of inference).

• Write down the meaning that you would impose on the cultural meaning (rung 3).

• Invent a possible solution to deal with such meanings.

• Write an actual conversation that produces the invention you just made.

Feel free to make all the changes you wish during the exercise. Every change is a sign of learning and another opportunity for practice. This is not a win/lose competitive situation with yourself or with others.

Again, put your written work away, this time for at least a day, before rereading it. If you prefer not to wait for the week, or the day, show your efforts to someone else. It is best to do this with persons who are also interested in learning about themselves, and who might reciprocate by showing you a case they had written in this format. The set is then one in which both of you are learning.

Another step is to make exercises like those just described part of the firm's executive-development activities. For example, you and your group members could each write a case. One of these could be the subject of discussion during a seminar. (I recommend at least an hour and a half and some trained professional help for each of the early sessions.)

Another possibility is for all individuals to write a case about an organizational problem that plagued, or continues to plague, the organization. It is then possible not only to see how each player conceptualizes the problem, but also how he or she has tried (or would try) to solve it.

During these discussions, the players soon generate a lot of data about the organization, its culture, and the way decisions are made. Grouping the cases provides a new data source for diagnosing organizational features that inhibit or facilitate orga-

nizational learning. In a large professional firm, for example, top management realized that if the partners were going to be successful in dealing with mediocre performance, they would have to become much more candid and forthright. They also realized that such candor was countercultural; hence the partners had to learn new skills. They also realized that if the partners learned to confront constructively, and if it worked well with the subordinate professionals, the latter would probably take them up on their challenge and start confronting the partners and the firm's policies.

Top management believed this would lead to constructive dialogue and possibly a new culture. All the top executives agreed with this espoused policy. However, as they examined their scenarios and the self-initiated censorship (what they placed on the left-hand side of their page), they became aware that their theory-in-use was quite different. About half of them were trying to act in ways that were consistent with the new policy. About half were easing-in but denying this was the case. One senior executive then said:

> Let's assume for the moment that our subordinates will be watching not only what we espouse but how we act. If that is so, then many of them should be aware of the bipolar nature of our actions. Yet to my knowledge, this is never discussed. They give much lip service to our policies. This means in the name of candor we must be (unknowingly) helping them to identify what is undiscussable, to keep it undiscussable, and to act as if they are not doing so!

The next phase is to use the new skills in everyday situations. In one firm, for example, professionals at all levels went through the X-Y experience. When the officers had to evaluate their professionals in the normal review process, they decided to use the opportunity to practice their new skills. Often they needed practice sessions ahead of time to help them prepare to conduct the review appropriately. The number

and length of the preparatory sessions were greatly reduced once the officers felt secure in their new skills.

Since the bewilderment, bafflement, and frustration of the X-Y case were experienced by all, subordinates who came to the review sessions knew how difficult it would be for the officer to behave in line with the new model. It was easy for the officers to say that the evaluation was going to be a learning experience for them as well as for the subordinates. If the subordinate agreed, I participated in the evaluation sessions. The subordinate and the superior often listened to the tape of the session, reporting that doing so was an eye-opener for them.

Another type of intervention is illustrated by the officer who asked his project team to reflect on a recent job they had completed. Although everyone involved evaluated the project as a success, he felt that they could have done better since they were exceptionally talented professionals. In the first session, team members reflected on their experience of the project. They identified ten factors that led the team to be less creative than it could have been. For example, they admitted that, as senior professionals, they had all acted as chiefs, each in his own bailiwick, adding that they had often recognized a lack of coordination and integration throughout the project. Their conclusion was that the team could have gained by stronger leadership.

A second session was held to dig into the reasoning that the team used during the first meeting. I noted, for example, that the team had recognized the lack of coordination early in the project. "Do you recall what prevented you from surfacing these views?" Their replies clustered around (1) the fear that they would step on someone's toes, (2) the fact that they were all very busy, and (3) the assumption that team coordination was the officer's responsibility. The officer 17

mentioned that he had reduced his time with the team because top management had asked him to take on additional and unexpected work. He had agreed to the extra work in part because the team was so senior that he believed they could administer themselves.

I asked what cues or data the members got that led them to believe that they would be stepping on other people's toes. The responses ranged from cues individuals gave each other to informal policies of the firm that made them hesitate to speak out. Their theory-in-use was to cover up their views; to cover up the cover-up in order to get the job done. This deeper analysis pointed to a different change target. If they could change the need for the cover-up, if they could learn to generate more effective cooperative relationships (including being able to make the heretofore undiscussable issues discussable), then the need for a strong leader might be greatly reduced. If so, the firm could use its senior officers more flexibly.

An episode in another firm involved design of a new organizational structure. Sides were taken; subgroups began to view each other as conservatives and liberals. In X-Y case terminology, each side held untested and often unillustrated attributions that led to divisive intergroup dynamics. Instead of reenacting their history of intergroup warfare, several executives pointed out the connection between the reasoning they used in the X-Y case and the reasoning they were currently using. A meeting was held with the key players on both sides present. In presenting their analysis of the situation, several executives asked, "If this makes sense to you, would you all be willing to join us in reflecting on what we are doing?" Most agreed spontaneously; a few agreed but were concerned—As one commented, "This could lead to blows." Well, it never did. They were able to map the attributions and evaluations
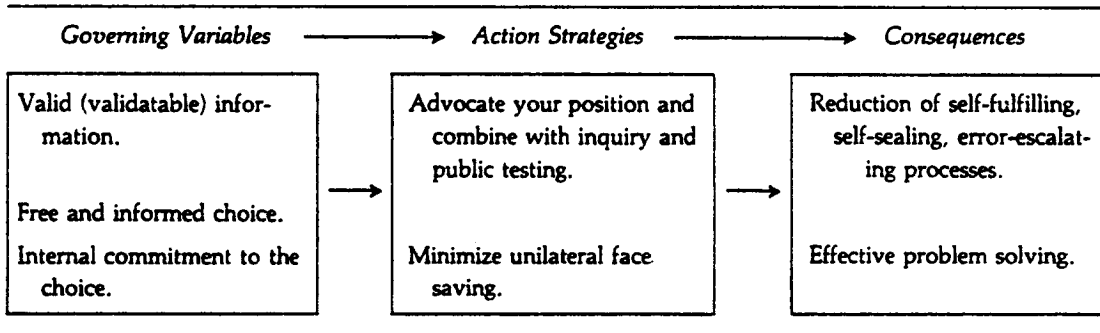
people were making and not testing, the games that were being played, and the possible negative consequences of all this on their final decisions and on the firm. The result was a jointly developed plan on how to reduce counterproductive factors. The participants agreed that the result was a greater degree of internal commitment to make the new design work, as evidenced by their willingness to monitor it actively and to design and implement changes that would make it more effective.

One of the other important consequences of the exercise was not learned until several months later. Many of the executives at the middle and near-top level—especially those who tended to play it safe—were baffled by the degree of commitment, and especially by the near absence of undiscussable issues, on the part of the top people. Many had predicted that it would take years to implement the plan. They told their subordinates not to become too anxious because top management would probably be changed before the new scheme had an important impact on their levels. This led middle-level executives to reduce their vigilance and concern about implementation. However, unlike previous occasions, the implementation not only moved faster but greatly reduced the space for hiding. This taught the middle and lower levels a more vivid lesson about the change in the firm than could have occurred in the usual information and exhortation exchanges.

Model II Theory-in-Use

Embedded in the advice above is a different theory-in-use that we shall call Model II (Figure 5). Its governing variables are valid information, free and informed choice, and an internal commitment to that choice in order to monitor the effectiveness of the implementation of the action. Model II is a theory

Figure 5

MODEL II THEORY-IN-USE

Governing Variables ⟶	Action Strategies ⟶	Consequences
Valid (validatable) information. Free and informed choice. Internal commitment to the choice.	Advocate your position and combine with inquiry and public testing. Minimize unilateral face saving.	Reduction of self-fulfilling, self-sealing, error-escalating processes. Effective problem solving.

of action that combines learning and inquiry with advocating one's views. It is not a nondirective model (such a model would simply be the opposite of Model I). Model II action strategies are to combine advocacy with inquiry, to minimize face saving, and to encourage the acceptance of personal responsibility.

Let us now consider the X-Y case with a view to redesigning it along lines more consistent with Model II. Recall that the meanings the executives produced were high up on the ladder of inference, unillustrated, and untested; that executives believed their evaluations and attributions were *low* on the ladder of inference — that is, they were concrete and obvious inferences — and that it was therefore unnecessary to test them.

An outstanding feature of the participants' evaluations of Y was their negativeness. The most probable explanation for this negativeness was that Y's actions *were* negative. What is ineffective about communicating negative judgments if they are valid? The answer, from this perspective, is "Nothing." The problem is that negative evaluations should not be communicated by using the same features that the actors believe it is ineffective for someone else to use. To the extent that meanings are communicated in a way that follows the same causal theory the sender has told the recipient is counterpro-

ductive, the sender will be experienced by the recipient as behaving inconsistently and unjustly.

The participants' attributions had two major features. First, they explained Y's actions by attributing motives "in" Y. For example, Y was protecting himself, was seeking to frighten X, intended to intimidate X, and was insecure. Second, the attributions stated that the cause was in the role or the position that Y held. For example, Y was acting like a company man, like an authoritarian boss, like a superior identifying with the hierarchy.

Not only do the attributions contain negative evaluations — they imply that Y intended to make X defensive in order to protect himself or the organization.

If we combine the features of the evaluations and attributions just described, and if we keep in mind that the receiver is also programmed with Model I, then we have the basis for the predictions of self-fulfilling prophecies, self-sealing processes, and escalating error that were illustrated.

To invent and produce a different way of dealing with X, let us first identify the key features of an intervention that contains negative or threatening meanings even through the intention is to facilitate learning.

Messages should be designed so that they are experienced as credible by the 19

recipients. Recipients must have access to the data and to the reasoning used by the sender to arrive at his or her evaluation or attribution. Hence the evaluations or attributions should be illustrated and the reasoning made explicit.

The message should be communicated in ways that will minimize individuals' automatic response to defend themselves. This means that senders should state their messages in such a way as to encourage inquiry into or confrontation of their reasoning and meanings.

From the preceding, it is possible to infer rules for producing such messages:

1. Provide the (relatively) directly observable data (first rung on the ladder) that you use to infer your evaluations or attributions, and check to see whether the recipient agrees with your data.

2. Make explicit the cultural meanings that you inferred from the data and seek confirmation from the other person.

3. Make explicit your judgments and opinions in ways that permit you to show why the consequences of the actor's action were inevitable, but without implying intentions to produce such consequences.

4. Encourage others to express feelings or ideas that they may have about the process.

The reader might ask, "How efficient can such rules be? Can we get anything done under real-time constraints?" First, recall that these rules are for dealing with double-loop issues. Second, how much is actually accomplished in the present modes of communication? Research suggests that present modes actually take longer and, worse, that they generate a social pollution of misunderstanding and mistrust that gives people a sense of helpless hopelessness. As our world becomes saturated with this pollution, even a small incremental error can touch off a breakdown of our natural defenses against it. As we shall see, the process of designing and implementing mean-

ings in accordance with Model II does not take much longer. When we have clocked Model I and Model II roleplay, we have discovered that Model II (when produced by actors who are moderately competent) usually requires the same amount of time as the Model I roleplay, or even less.

We now turn to an illustration of how the interventionist might deal with Y. Remember the ladder of inference and Y's statement that led us to infer the meaning "X, your performance is unacceptable." Remember, too, that inferences are subject to error and hence should be put to public test. Every move up or down the ladder necessarily means that inferences are made; the higher up the ladder of inference and the more abstract the ideas, the greater the chance of error and therefore the greater the importance of public testing. Whatever theory one uses, it should make public testing as easy as possible.

To test an inference with someone else, it is necessary to make explicit both the premise and the conclusions drown from the premise. The inference "Your performance is unacceptable" in the X-Y case is based on the premise of Y's words to X.

One can test the inference by asking Y: "When you said, 'X, your performance is not up to standard,' did you mean that his performance was unacceptable," or "When you said 'You seem to be carrying a chip on your shoulder,' did you intend to attribute to him unacceptable attitudes?" If Y responds yes, then the meanings have been affirmed and it is possible to proceed to the next rung on the ladder.

If a participant in an X-Y seminar wanted to reveal his diagnostic frame, it is at this point that he would have to say something like "Well, I infer from these data that you prejudged X," or "You were too blunt," or "You were insensitive." Such a response is likely to produce defensiveness in Y for several reasons. First, Y may not agree

with the evaluations/attributions. Y may believe that he had to be blunt or insensitive in order to get through to X. Or he may believe that he did not prejudge X, that X generated years of data that led to the present judgment. Second, the evaluations not only attribute errors to Y, but imply that he intended to be blunt, insensitive and so on. Since no one knowingly produces error, if Y knew what he was doing, then he knew that he was being blunt and insensitive. An explicit negative evaluation is coupled with an implicit attribution that Y intended the encounter with X to produce these negative consequences.

Under these conditions, the interventionist's testing is more of a trick than a helpful strategy. He or she may have tested the first two levels of inference only in order to nail Y with his or her third-level evaluations and attributions, themselves difficult to test. Indeed, in our experience it is difficult both for the interventionist and for the participants to see and agree with the logic of inference between successive levels of inference. Recall how often in the transcripts the participants either were unable to illustrate their inferences or illustrated them with further inferences. If inferences are to be tested, then no matter how high on the ladder of inference they occur it should be possible to proceed down the ladder and explicitly connect them with the first and second levels.

To summarize: Whatever concepts are being used, one should be able to order them on a ladder of inference, advancing from relatively directly observable data to the culturally acceptable meaning and then up to the concepts used to organize the previous two rungs into a problem. It is at this point that interventionists are introducing their own (usually tacit) theory of help.

Let's return to the interventionist's two questions to Y. Recall that Y confirmed the meanings. But let us assume that Y said *yes* but was showing signs of impatience: "Of

course I meant the performance is unacceptable!" or "Naturally I think X's attitudes are wrong! What are you driving at?" At this point the interventionist, using our theory of action, could say:

I'll be glad to tell you what I am driving at. First, though, I want to make sure that I understood you correctly.

I have a way of understanding the effectiveness of the kinds of comments that you made to X. Your first comment (repeats it) I call an "unillustrated evaluation." It tells the person he is wrong, but it does not include the data and logic of how you arrived at that conclusion.

People tend to react to these unillustrated evaluations and attributions by feeling bewildered and/or misunderstood. Depending on how free they feel, they may confront you or they may imitate your style and make their own unillustrated evaluations and attributions about you. If they do the latter, it upsets the receiver, just as X was upset by your comments. Now, if X reacted on inferences that he is keeping secret, you would probably sense that secrecy because you would not see clearly the reasoning he used to come to his stated conclusions.

Let's stop for a moment. What is your reaction? Does this make sense? (or) Am I communicating?

Several features of this response should be highlighted. First, not only do the concepts of "unillustrated evaluation" and "unillustrated attribution" provide insight into a problem, but the insight is in the form of a causal theory. *If someone produces unillustrated evaluations, whoever receives the evaluations or attributions will not know the basis of them unless they are illustrated. The receiver will feel bewildered and misunderstood. He may therefore react defensively (unless he is afraid or prefers to be dependent on you).*

The causal theory in these propositions is true for anyone, not simply for Y. Therefore, framing it as we do gives Y a degree of distancing from the problem that may help him or her to understand it better. We are not saying "Y, you are wrong." We are saying that anyone who behaves as Y did 21

will produce the unintended consequences described above.

When individuals observe the redesign of what they could have said to Y, they are often impressed with its simplicity and obviousness. Many report that they expected the answer to be more surprising. The redesign may not be surprising because the ideas behind it are self-evident and not new. Another reason is that many individuals report that they had considered some of these intervention ideas but did not know how to design responses that made sense.

Both of these reasons show how crucial is the distinction between being aware of a possible action and being able to produce that action. In our terms the roleplay may not seem surprising, because it fits many individuals' espoused theories. But the difficulty is evident when people try to produce such redesign and find themselves unable to do so. They are surprised that they have difficulty redesigning their own interventions. The reason for their difficulty is that they *still hold* a Model I theory-in-use. When they are listening or advising, individuals use their espoused Model II theories. When they try to produce action, their Model I theory-in-use is activated.

The governing values of the theory are especially important. It is possible to make Model II statements to Y and yet fail because the actor still holds, and subtly conveys, Model I governing values such as "Win, don't lose," or "Maintain unilateral control." Indeed, some people initially react to the Model II redesign as if they were utilizing a new and more subtle form of Model I. Their disguised Model I approach usually shows when they try to defend their views.

To close, the skills and competencies that executives learn for dealing with an X-Y type of problem can be used for dealing with any double-loop problem. The key is to learn the new skills *and* to acquire a new set of governing values. If executives learn the new

skills — such as advocating their position and encouraging inquiry — but use them to maintain unilateral control and to maximize winning, they would be using their new skills in the service of Model I values. They remain within a Model I mode; they hide their views about the gimmickiness of the new behavior, yet act as if they are not hiding anything. As a result, others may interpret their newly acquired skills as gimmicks or as new ways to manipulate people.

Luckily, people judge the credibility of human skills by evaluating what values they serve. This means that those who learn the new skills as gimmicks and tricks will be discovered. It means further that those who wish to gain credibility not only must learn the new skills, but also must internalize a new set of values.

ⓓ

SELECTED BIBLIOGRAPHY

The ideas in this paper come from the author's *Reasoning, Learning and Action: Individual and Organizational (Jossey-Bass, 1982)*. Also relevant is a study of a group of presidents learning Model II, *Increasing Leadership Effectiveness* (Wiley-Interscience, 1976). Shorter versions of the theory behind the research may be found in the author's previous *Organizational Dynamics* article, "Leadership, Learning, and Changing the Status Quo" (Winter, 1976).

Two studies in government policy making and regulation are Laurence E. Lynn and David de F. Whitman's *The President as Policymaker: Jimmy Carter and Welfare Reform* (Temple University Press, 1981) and Robert B. Reich's "Regulation by Confrontation or Negotiation" (*Harvard Business Review*, May-June 1981).

The research on managing paradoxes can be found in Louis B. Barnes's "Managing the Paradox of Organizational Trust" (*Harvard Business Review*, March-April 1981).