Resource papers in action research

Argyris and Schön:
some elements of their models [1]

Bob Dick and Tim Dalmau

This is a resource file which supports the regular public program "areol" (action research and evaluation on line) offered twice a year beginning in mid-February and mid-July. For details email Bob Dick bdick@scu.edu.au or bd@uq.net.au

... in which some of the elements of the models of Chris Argyris and Donald Schön are briefly described

In this paper ...

-

-
-

As the Johari window clearly implies, much of what goes through our mind is not expressed. In some instances, perhaps this is just as well &endash; would you care for your every thought to be made public? But some of what we usually hold secret can be used differently to advantage. It can improve our own effectiveness, enhance the quality of the relationships we enter into, and renew the organisations and social systems we inhabit.

This first part describes some concepts which can help bring about the improvement.

We begin with an overview of the notion of "theories of action" &endash; that certain assumptions about the nature of the world underlie much of what we do. We then examine in turn the information which is relevant to understanding interaction and other behaviour, the sets of assumptions which guide us in our behaviour, and the effects of different world views on learning. Together, these concepts support the tools, practical processes and techniques which form Parts 2 and 3 of this book.

Theories of action

Central to the work of Argyris and Schön is the concept of a *theory of action*. Theories of action are the "mechanisms" by which we link our thoughts with our actions.

They have a number of elements...

1   Action strategies

2   Consequences for self

3   Consequences for others

4   Governing values

5   Action strategy effectiveness.

Each of these is explained further below.

Argyris and Schön divide theories of action into two types...

1   espoused theories

2   theories-in-use

Espoused theories are those which we know about: which we espouse to ourselves. Theories-in-use are the theories of action implied by our behaviour; they are more likely to be unknown to us (Figure 2). *[Note: We've preserved the same figure numbers as in the book this is taken from.]*
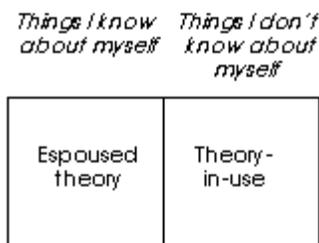
Things I know about myself	Things I don't know about myself
Espoused theory	Theory-in-use

Fig. 2

Espoused theory and theory in use related to the Johari window

Action strategies: These are the behaviours in which we engage to manage our immediate surroundings, especially our social surroundings. Argyris would say that they are to keep a governing value within an acceptable range: to maintain an important belief.

Consequences for self: These are the end effects for ourselves of our action strategy and of the response it engenders in others. It often includes what we feel obliged to do or prevented from doing.

Consequences for others: These are the end effects for others of our action strategy and the response it engenders in them; often they include what they feel obliged to do or prevented from doing. "Others" can include people, groups, organisations or systems.

Governing values: Governing values or governing variables are constancies which we seek to keep within acceptable ranges. They are goals we seek to satisfy, beliefs we seek to operationalise or defend, values we seek to express... For example...

- to maximise winning and minimise losing;

- to minimise expressing negative feelings;

- to be rational;

- to decrease the opportunity for confrontation;

- to define the group task unilaterally and have the others agree to it;

- to maximise cooperation and collaboration;

- to maximise free and informed choice.

In the book *Reasoning, learning and action* Argyris uses the term *governing value* interchangeably with *governing variable* in a number of places. In our view, a governing variable or governing value is best thought of as a mix of motives, values, beliefs and feelings, the specific mix depending on the person, the situation and the context. We will mostly use the term *governing value* as more nearly self-explanatory.

Action strategy effectiveness: This denotes the extent to which our behaviours (our action strategies) lead us to confirm the "rightness for us" of our governing values. The effectiveness of our action strategy is judged in relation to the governing values of either our espoused theory or our theory-in-use.

An example...

> *"For one person, the value of winning may be very important (Governing [value]) ... I will hesitate to cooperate with others if this means that they might win instead of me (Action strategy) ... The person bent on winning is viewed with suspicion. The legitimate needs of other people are pushed aside and the goals of the organisation become secondary to winning (Consequences for self and others) if this can be done without blowing his or her cover"* [Egan: 1983, p.XIV-5].

These general concepts can now be transformed and modified slightly to render a working model for interpersonal and intergroup behaviour. This will provide much of the specific basis for the processes which are described later as tools and applications. We will use terminology that is probably a little more familiar to the reader.

To develop this working model, we analyse the nature of relationships. We examine the types of information which may be needed to resolve difficulties in relationships. The model which results will be found to guide many of the applications which form the later parts of this book.

Subsequent sections then further examine Argyris' concepts in the light of this material.

[index]

Relationships and information

Some relationships are effective and satisfying. When the occasional problems arise they are quickly and easily dealt with. If one person indicates dissatisfaction, the other cooperates in identifying and remedying its source.

Other relationships, even those that may be effective and satisfying most of the time, have a different experience. From time to time problems arise without being dealt with effectively.

Consider the types of information required to resolve different problems. As an example, it may be that the problems are...

- recent,

- work-related, and

- with not much feeling invested in the issue.

If so, exchanging relevant information about work requirements may be enough to remedy the problem.

In other words, the problems can most likely be resolved using role-based procedures which focus on "objective" information. In terms of the model we are developing here, the required information consists of the following...

$$\text{actions} \quad \genfrac{}{}{0pt}{}{\rightarrow \text{ material outcomes} \rightarrow}{\leftarrow \text{ material outcomes} \leftarrow} \quad \text{actions}$$

One person's actions produce material outcomes for the other person. In turn, the second person's reactions produce material outcomes for the first. The result is often the cycle depicted immediately above: how each person reacts to the other maintains the present relationship.

Depending on whether the outcomes are harmful or useful, this is a simple form of a vicious or

beneficial cycle. Most such cycles are more complex. Often, there are psychological as well as material outcomes resulting from each person's behaviour.

So, often, issues arouse strong feelings (whether those issues begin as problems within the work relationship or not). Information about feelings and about the relationship may then be needed to bring about a resolution. Suitable procedures might include relationship-based team-building, or conflict resolution (which is very similar in many respects). The information exchanged will include information about feelings, and the actions and other specific consequences of those feelings...

$$\text{actions} \begin{array}{l} \rightarrow \text{outcomes} \rightarrow \text{feelings} \rightarrow \\ \leftarrow \text{feelings} \leftarrow \text{outcomes} \leftarrow \end{array} \text{actions}$$

The types of relationship problems discussed so far are relatively simple. You will know from your experience that some relationship problems are more complex and long-standing. It is then usual for people's beliefs about each other to be an important part of the issue.

Conventional teambuilding and conflict resolution procedures, used in such circumstances, are frequently only partly effective. They may at first appear to provide a resolution; but often the relationship breaks down again. When this occurs, it is often because of the "hidden agendas" which exist. To resolve the issue in a more lasting way may require surfacing the beliefs people hold about each other. This is where the contributions of Argyris and Schön become particularly valuable.

The overall form of information to be exchanged now includes actions, outcomes, beliefs, and feelings...

$$\text{actions} \begin{array}{l} \rightarrow \text{outcomes} \rightarrow \text{beliefs} \rightarrow \text{feelings} \rightarrow \\ \leftarrow \text{feelings} \leftarrow \text{beliefs} \leftarrow \text{outcomes} \leftarrow \end{array} \text{actions}$$

People are often very reluctant to describe with honesty their beliefs about and feelings toward each other (or for that matter, about themselves). In a difficult relationship, therefore, each person is denied some of the information needed to understand the other person's actions. The misunderstanding therefore persists.

Currently, many of the procedures for building relationships don't exchange information about beliefs. Here, we think, lies part of the importance of the contribution that Argyris and Schön have made. [2] They have recognised that beliefs, particularly about other people's motives, give rise to some of the more harmful and less effective dynamics of relationships. They argue (successfully we believe) that when we act on unspoken beliefs about others, harmful ways of relating to each other are created.

To our knowledge, Argyris seldom (if ever) discusses categories of effective information in precisely these terms. However, if you compare the categories of information to his verbatim accounts, you will find that there is a close correspondence. The use of these categories of information can therefore help people acquire a less defensive communication style. They can also be an aid to the design of processes, as illustrated by the later examples.

Notice, too, that the elements of information are linked together. Actions by one person produce material outcomes for another. The second person then develops assumptions (beliefs) about the first person, and these influence her attitude (feelings) towards that person. If she then *acts out* her feelings, her actions may well trigger a similar sequence in the opposite direction. In the remainder of this book, we will talk about this dual sequence (or each half of it) ...

$$\text{actions} \rightarrow \text{outcomes} \rightarrow \text{beliefs} \rightarrow \text{feelings} \rightarrow \text{actions}$$
$$\text{actions} \leftarrow \text{feelings} \leftarrow \text{beliefs} \leftarrow \text{outcomes} \leftarrow \text{actions}$$

as an "information chain".

[index]

The information chain

A further important feature of this information chain is that it may allow a mutual self-fulfilling prophecy to emerge. In a difficult relationship it is common for each person to develop suspicions about the other person's motives.

Seldom if ever checked rigorously, these suspicions may strengthen over time. They may come to be treated as "facts". Each person may then experience her own behaviour as a *reaction* to the other person; yet that very reaction is what allows the other person to feel justified in her stance.

Before moving on, we want to draw your attention to one feature of this information chain. We also intend to add another element to it.

The feature is this. Some of the elements of information are available to one person, some to the other. Typically, no element is available to both. Therefore, typically, both people are able to maintain a different perception of the situation. Because there are social rules against expressing some of this information, it typically remains hidden.

In our initial use of the Argyris and Schön models, we made use of the assumption that different people have access to differing information. For illustration, imagine that two people, A and B, have an ongoing relationship. Then, typically, the information each of them *can* access under some circumstances is as follows...

A's information	B's information
B's actions	A's actions
material outcomes for A	material outcomes for B
A's beliefs about B	B's beliefs about A
A's feelings towards B	B's feelings towards A

We would add the important proviso that people often have poor access even to their *own* beliefs and feelings.

In the past we have treated actions as to some extent accessible to both people. Actions and material outcomes we treated as at least verifiable. We saw actions and material outcomes as offering a common language through which agreement could be generated. In turn, beliefs and feelings can become verifiable to some extent by being linked to the more tangible actions and outcomes.

We still believe this can be so. In fact, we make use of it in some of our processes. In practice, however, we sometimes find surprising disagreement about what a person had done. A's perceptions of her behaviour are often at some variance with B's perceptions of the same behaviour.

Further exploration and thought has revealed some reasons for this. First, people are not very effective observers of their own behaviour. (Our eyes are placed for observing others, not ourselves). Secondly, people tend to judge their own behaviour by their intentions. They judge other people's behaviour by its outcomes.

We therefore sometimes add to the information chain a further element: intentions ... [3]

$$actions \begin{array}{c} \rightarrow outcomes \rightarrow beliefs \rightarrow feelings \rightarrow intentions \rightarrow \\ \leftarrow intentions \leftarrow feelings \leftarrow beliefs \leftarrow outcome \leftarrow \end{array} actions$$
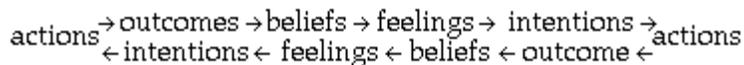
This allows for people often being unable to act out their intentions because of situational constraints, for example, or a lack of the necessary skills. It also increases the model's correspondence which that of other writers. Miller and his colleagues (1979) include *intentions* amongst their five categories of information. Fishbein and Ajzen (1975) also use as their core model something very similar to later part of the information chain...

$$belief \rightarrow attitude \rightarrow intention \rightarrow behaviour$$
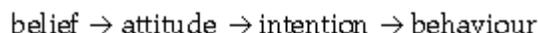
The information chain, you will find, is easy to convert into action. Equipped with this practical tool, let us now return to the ideas of Argyris and Schön.

[index]

The information chain
and the theory of action

It is time to draw a parallel between these more interpersonal aspects of behaviour and the general Argyris and Schön model of a theory of action.

To relate the information chain to Argyris' concepts, consider a person's initial actions as equivalent to an action strategy with consequences for self and for others. The outcomes are the consequences. Beliefs (and feelings) form a large part of a governing value of the theory-in-use. Intentions emerge from governing value of the espoused theory. Figure 3 shows this.

actions	outcomes	beliefs ➤ feelings	intentions
action strategy	consequences for self and others	governing values (in use)	governing values (espoused)
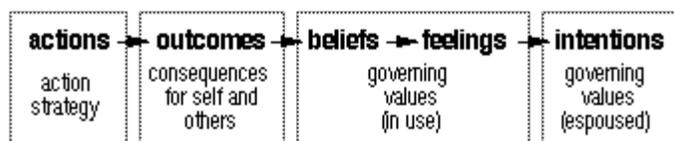
Fig. 3. Argyris' concepts and the information chain

More thoroughly, we can reveal the correspondence by teasing out two complete strands in the information chain. One of them relates to espoused values and strategies, and the other to theory-in-use and strategy-in-use. Figure 4 shows this more complete version. (For most purposes, the summary version in Figure 3 is adequate.)

intentions	outcomes	beliefs ➤ feelings	intentions
action strategy (espoused)	consequences self & others (intended)	governing values (espoused)	action strategy (espoused)
actions	**outcomes**	**beliefs ➤ feelings**	**actions**
action strategy (in use)	consequences self & others (actual)	governing values (in use)	action strategy (in use)
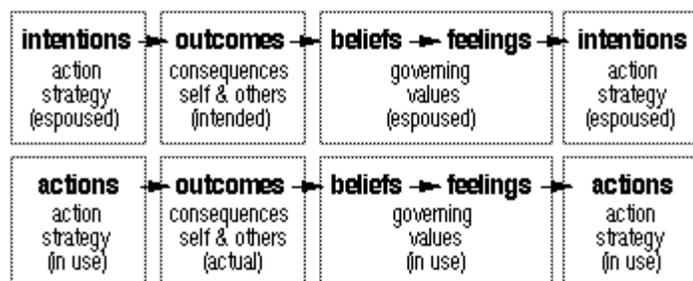
Fig. 4. Argyris and the expanded information chain

A person's governing values are typically expressed in an action strategy, which has consequences. Most commonly, the beliefs are assumptions about the other person's motives: unchecked ascriptions of the governing values of her theory-in-use. Behind these lies a world view, including a set of social rules governing behaviour.

Before proceeding, notice that communication depends upon a very highly developed and practised set of skills. Notice, too, that like many complex skills, most of it happens outside awareness. As we develop through our very early years, we absorb the complex communication style of those around us.

Michael Argyle and his colleagues (for example, Argyle and others, 1981) would say that we learn a complex set of social rules that we use fairly consistently, but often find hard to express explicitly. We can expect that when we communicate we do not always consciously know what we do. Conditions therefore favour the development of incongruence between theory-in-use and espoused theory.

Argyris and Schön suggest that we all have a strong propensity to hold inconsistent thoughts and actions. The links between what we think we are trying to achieve and the way we go about it are often not what we imagine: our espoused theories differ from our theories-in-use.

To put it simply, we *don't* always practise what we preach, however sincerely. The difference between espoused theories and theories-in-use applies at the level of national strategies, organisational management strategies, and small group and interpersonal behaviours.

Thus are we often less than effective in many of our behaviours. To improve this, we can achieve a better understanding of the links (Figure 5) between what we think we are trying to achieve and what we actually do. We then have more options for increasing our effectiveness and satisfaction, and also that of others.

Things I know about myself	Things I don't know about myself
Action strategy Consequences for self Consequences for others Governing values Effectiveness	Action strategy Consequences for self Consequences for others Governing values Effectiveness
ESPOUSED THEORY	THEORY-IN-USE

Fig. 5

Elements of the theory of action occur for both espoused theory and theory-in-use

When a person realises a mismatch between ideal-self and actual-self, strong feelings can be generated. This was brought home to one of us some years ago. One of us was conducting leadership workshops for school principals with Bernie Neville of La Trobe University.

The principals were given the task of generating a picture of ineffectiveness. They were asked first to describe a common failing of principals, one which they themselves would also "own up" to committing occasionally. They were asked to describe the behaviour, its consequences for them, its consequences for the other parties involved, and its overall effectiveness. They were then invited to imagine what appeared to be the overall value or belief which underpinned this strategy.

It was then suggested to them that the governing values often appeared to be some mix of the following...

- to maximise winning and minimise losing;

- to minimise expressing negative feelings;

- to be rational;

- to decrease the opportunity for honest confrontation;

- and to define the group task unilaterally and have the others agree to it.

A near riot developed. It seemed that they were prepared to discuss and analyse their own behaviour. But when it came to acknowledging that their "intentions" or "unconscious motives" were not those which they espoused, they became intensely resistant.

This miniature case study exemplifies much of what Argyris has often described and illustrated about the relationship between espoused theory and theory-in-use. For each of us...

- there is a gap between what we think we believe, and the values implied by our behaviour;

- we are blind to this gap;

- though others may perceive it, they are reluctant to admit that they have, let alone bring it to our attention;

- if they do, we are likely to react most defensively.

For good measure, these taboos against being open about beliefs and feelings are then reinforced by a taboo against revealing the taboo. Argyris (for example 1985) calls it the undiscussability of the undiscussable -- the cover-up of the cover-up.

This same inconsistency between espoused and actual values and strategies lies at the core of many difficulties in relationships between individuals and groups. In addition, people tend to ascribe incorrect governing values to others' behaviour. This is as true of individuals as it is of groups or, need we say it, nations.

At this point we can extend the Johari window. Our intention here is to accommodate these inaccurate perceptions that people have. What people think they know about each other may or may not be accurate. Figure 6 makes the distinction.

A core assumption, as Argyris and Schön have said, is that people seldom reveal their assumptions about each other, especially about motives. When they act on their assumptions, their motives are very often misunderstood. The common result is the mutual self-fulfilling prophecy we have already mentioned: each person's assumptions are maintained by the other's behaviour and support the person's own behaviour.

Things I know
about myself

Things others
think they know
about me, and
which are valid

| Action strategy |
| Consequences for self |
| Consequences for others |
| Governing value |
| Effectiveness |

Things others
think they know
about me, and
which are
invalid

| Action strategy |
| Consequences for self |
| Consequences for others |
| Governing value |
| Effectiveness |
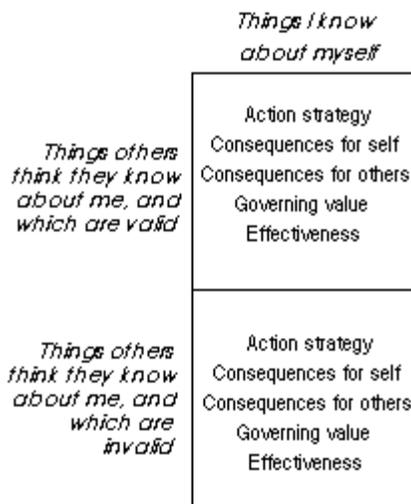
Fig. 6

What people
assume they
know about
someone may or
may not be valid

You can think of this as happening in two stages. At first, because the person's motives are unknown, her behaviour is different from what others expect.

At this point, something curious happens. People may not even *consider* that their expectations may be incorrect. They explain the discrepancy by attributing motives (as in Figure 7). You have probably heard someone say...

- *"She knows what she is supposed to be doing! Why isn't she doing it? She must be up to something."*

The second stage occurs when people act on assumptions. Because of the social taboos they usually don't voice them. They merely act on them as if they were true. This often results in more of the very behaviour which triggered the assumptions in the first place. The assumptions, people believe, have been confirmed. This is the self-fulfilling prophecy in operation.

I expect certain
behaviour

↓

I observe other
behaviour

↓

I explain the
discrepancy
by attributing
motives
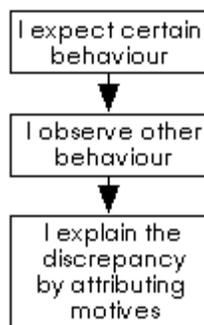
Fig. 7

We usually explain
unexpected behaviour not
by changing our
expectations, but by
attributing motives to the
other person

The actor's motivations are unknown to the other person, who may also thus attribute motives. A *mutual* self-fulfilling prophecy (Figure 8) may then result.
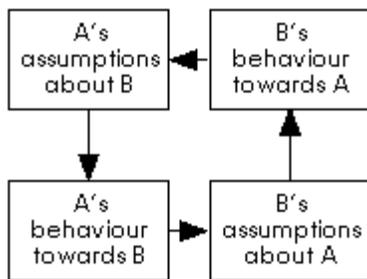
Fig. 8

A mutual self-fulfilling prophecy based on mistaken attributed assumptions about motives

Figure 9 shows a more detailed version of this.

The purpose of many of our applications of Argyris and Schön's work is two-fold. Firstly, it is to enable people to make contact with their assumptions about each other (the often incorrectly ascribed governing values, or perhaps the strategies which they assume drive the other person's behaviour). Secondly, the applications enable people to exchange this information in such a way that it can be understood, and challenged, and possibly corrected.

However, you cannot simply *ask* a person what her theory-in-use is. By definition, it is largely beyond her awareness. It has to be inferred (through successive approximations) from an examination of her behaviour, its consequences and the assumptions on which the behaviour is based. In terms of the expanded Johari window, it is a matter of encouraging people to compare their espoused theory (and all that goes with it) with their theory-in-use as revealed by others. Figure 10 gives a graphical summary of the comparison.
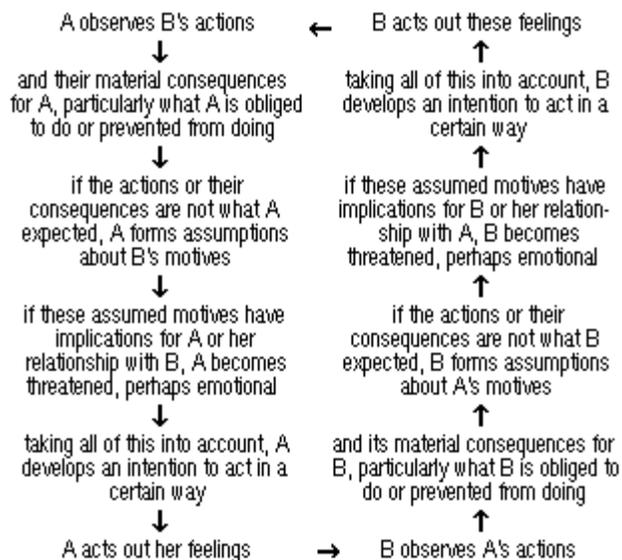
Fig 9. The information chain as a cycle

This is not a simple task: inferring a theory-in-use is a bit like inferring the principles of grammar from observed speech: devising progressively more adequate constructions. However, it is well worth it. Like climbing a mountain, the trek may be arduous but the view from the top makes the effort worthwhile.

Fig. 10.

Differences between what we think we know and others can reveal

Nor can you *observe* a theory of action directly, whether espoused or acted on. It can be inferred from observable behaviour, of which a person herself may only be partly aware, and which others can report.

Unfortunately, as we have already implied, there are strong taboos against telling people our beliefs about them and our feelings towards them. Assumptions about other people's motives are very seldom revealed. Even then they may be given in such a way that they are dismissed.

When relationships continue for any time, misperceptions develop. The actions which result often trigger reactions which (falsely) confirm the misperceptions. As we have said, this can be how self-fulfilling prophecies arise.

The misattributions can be corrected. In general this will occur only when the relevant information is exchanged and understood: information about attributions and the evidence for them. This is what many of the processes in Part 3 set out to do. Information is exchanged about the other person's actions, and these consequences for the speaker: material outcomes, assumptions (beliefs) developed, feelings, perhaps intentions, and the reaction.

Moreover, to analyse the links between behaviour and the reasoning processes on which it is based, it is crucial that you are able to describe *actual* behaviour and not behaviour that you *infer* of yourself or others. It is therefore important to have ways and means of concentrating on what...

- people actually did (or failed to do) and

- people actually said (or failed to say).

This point cannot be stressed too much. The processes outlined in Part II of this book are designed to do just this. However, we would also stress that this is not always as easy to do as it may first appear. We find the work of Michael Grinder [4] particularly helpful in this regard.

[index]

Models 1 and 2

We mentioned earlier a number of possible governing values, including "maximise winning and minimise losing" and "minimise expressing negative feelings". Some of these governing values tend to cluster together to form what might be called world views.

5

Argyris and Schön identified two such clusters. They term them Model 1 and Model 2. Model 1 might be described as a competitive and defensive stance towards the world. Model 2 is more collaborative and less defensive.

In more detail, Argyris describes the governing values of Model 1 as follows [1982, p.86]...

> *"Achieve the purposes as the actor perceives them*
>
> *Maximise winning and minimise losing*
>
> *Minimise eliciting negative feelings*
>
> *Be rational and minimise emotionality"*

This produces adversarial and defensive action strategies, poor relationships, and poor learning.

On the other hand, the governing values for Model 2 are [1982, p.102]...

> *"Valid information*
>
> *Free and informed choice*
>
> *Internal commitment to the choice and constant monitoring of the implementation"*

The action strategies here, less defensive and more collaborative, are also more conducive to effective relationships and learning.

In his work, Argyris has found that it is not unusual for people to advocate Model 2 values and to think they express them in action. That is, their espoused theories are Model 2. Our own view is that in our frequently adversarial culture people often show a mix of Models 1 and 2 in their espoused theories. To the extent that they think the situation allows them, however, many are willing to move as far towards Model 2 as possible. We agree with Argyris that their actions, especially when under threat, show the defensiveness of Model 1.

In his books and papers Argyris gives many verbatim accounts of the work he does. It is obvious from this material that people do often think they are acting out of Model 2 values. Whenever there is a potential for threat in the situation, however, they are likely to behave in the defensive ways which imply Model 1 values.

For example, he often asks them to give feedback to each other on their behaviour, or to role play the giving of feedback in case-study situations. Their feedback is given in an attacking way, or so carefully and tentatively that the information is "fuzzy" &endash; he describes the latter as "easing in". In the terms that we have used in other documentation (for example, Dick, 1986), they use the Model 1 actions of blame and criticism and demand, or talk in generalisations. They do not actually use the specific and non-defensive communication which they advocate to others and believe they are trying to use.

We have already talked about the way in which governing values are expressed in action strategies which in turn have consequences. We will shortly describe the way in which Model 1 values produces single-loop learning. Following Argyris, one might summarise the results as follows ...

Governing values	*Action strategy*	*Relationship consequences*	*Learning consequences*

Pursue own goals	Control the environment	Perceived defensiveness	Self-fulfilling processes
Play to win in win/lose style	Control the task	Defensive relationships	Single-loop learning
Minimise negative feelings	Unilaterally protect self	Defensive norms	Attributions are untested
Rationality, not emotionality	Unilaterally protect others	Low choice, commitment	Single-loop learning

Compare this summary to the conventional beliefs about proper behaviour in many organisations and social systems. You will probably agree with Argyris that most of our social systems are Model 1 systems.

We address some of the reasons for this in our monograph *Politics, conflict and culture*, especially the discussion of the social change model we call the second watershed.

To summarise the main features of the second watershed model ... The wider culture in the developed world is in many respects autocratic, and rational in the narrow sense of the term. Stable environments allow bureaucratic and controlled organisations to operate profitably within this culture. Substantial change (for example through double-loop learning) happens rarely and slowly. When environmental change is also slow, this may not be a serious problem.

An important feature of such systems is that there is a "right behaviour" for any situation. People are taught fixed solutions to cope with both their tasks and their relationships. The social rules they follow are detailed, specific, and often held as beliefs without reason.

In many organisations people are not encouraged to take responsibility for their actions. Superiors are responsible for subordinates' actions. Subordinates follow orders, as revealed by a reliance upon procedures and rules.

In one organisation where one of us was helping with cultural change there were over 70 procedure manuals of one form or another. People previously were expected to follow them carefully in their work. This organisation was trying to change into a more responsive system. As part of a change workshop a senior manager informed the rank-and-file participants that in future, procedure manuals were to be merely *guides* to action &endash; the employees were expected to work out for themselves what should be done, and to do it. The astonishment on participants' faces signalled this as a virtual revolution.

It is hardly surprising that in such a system the reasons behind behaviour are often lost in history. It is also to be expected that people assume that other people "know the rules", and know what they mean. Any perceived departure from the rules is then likely to be labelled disobedience even if it is ignorance or a differing interpretation.

At times such as this, when organisations face constantly changing environments, the difficulties multiply. The tacit rules are yesterday's rules. The more rapid the environmental changes, the more likely that the rules are inappropriate and the interpretations of them varied.

Organisations (and other social systems) display a number of other important features which add to this effect...

Power is one. When there are power differences between people, it is often tempting to "solve" problems (or at least dispose of them) by fiat rather than understanding.

Specialisation is another. People pursue the goals of their own position or section or branch, and may thus lose sight of the overall goals. This makes for competition rather than collaboration between person and person and between section and section.

Designing organisations around jobs rather than around people is a third. People are expected to leave at home the more individual aspects of themselves, including their thoughts and feelings. A false rationality results.

The result is that people are again denied the information they need to understand what is happening. Lacking this, they use competitive and adversarial methods if personal or sectional well-being is at risk. In many organisations and other social systems you don't get better resources by being honest. You do better by telling selective truths or, if you can get away with it, plausible lies.

Notice, too, that almost all of our most influential social institutions are adversarial in nature. Our courts of law, our political systems, our meeting procedures, our examination procedures, our industrial relations systems &endash; these and more are adversarial in important respects. People are *encouraged* to present information selectively. Those who are honest risk being treated as deviants.

To someone willing to observe closely and put aside preconceptions, our Model 1 social rules are demonstrated often.

Compare the description above to a Model 2 system. Model 2 is clearly a very different theory of action. It requires that people operate more consensually, with more openness about beliefs and feelings. It goes counter to many of the implicit rules we follow. This, we presume, is one of the reasons why cultural change (or personal change for that matter) is so often difficult.

Model 2 may be summarised as follows...

Governing values	Action strategy	Relationship consequences	Learning consequences
Valid information	Create environments which allow personal freedom	Minimal defensiveness	Disconfirmable processes
Free and informed choice	Joint control of the task	Non-defensive relationships	Double-loop learning
People responsible for own	Joint protection of self	Learning-oriented norms	Public testing of attributions

behaviour			
Openness	Bilateral protection of others	High choice, commitment	Double-loop learning

In other words, Model 1 processes encourage only superficial learning. The underlying assumptions are treated as fixed. Model 2 processes allow more substantial learning and change by allowing values and assumptions to be tested.

[index]

Learning processes

We have mentioned already two important concepts from Argyris and Schön. One is the notion of theories of action (espoused or in-use). A second is the two clusters of governing values, Model 1 and Model 2. A third cornerstone of their work is the concept of a learning cycle for organisations and groups. They suggest there are two basic types (pure types, we might add, and seen only in extreme cases). They tend to correspond with Model 1 and Model 2 systems.

Argyris says...

> *"People programmed with Model I theories of action produce Model I group and organizational dynamics that include quasi-resolution of conflict, uncertainty, avoidance, mistrust, conformity, face saving, intergroup rivalry, invalid information for important problems and valid information for unimportant problems, misperceptions, miscommunication, and parochial interests." [Argyris, 1985, p. 88.]*

Some readers might be forgiven for recognizing their own working or other social environment in this description. Indeed, as we pointed out above, Argyris suggests most of our social institutions are Model 1 systems. But how are such characteristics maintained? What is it about the *processes* that lead such systems to be self-replicating in their mediocrity and ineffectiveness?

Some of the reasons have already been implied in the previous discussion. In times of environmental stability, the systems which develop are hierarchical and conservative. People are given little responsibility for their own behaviour. They pursue their own goals mostly through adversarial processes.

Argyris (1985) provides the concept of a learning cycle to explain this maintenance in organisations and groups. He distinguishes those which grow, develop and adapt from those which might better be described by the quote above. He uses the terms single-loop and double-loop learning, attributing their use originally to Ashby (1952).

In its simplest form, single loop learning involves the generation of *new action strategies* to achieve *existing governing values*. Double-loop learning, on the other hand, involves adaptation and modification of the governing values themselves. As Argyris says...

> *"In single-loop learning, we learn to maintain the field of constancy by learning to design actions that satisfy existing governing values. In double-loop learning, we learn to change the field of constancy itself." [Argyris and Schön, 1974, p. 19].*

An example may illustrate the point... One of us has a cruise control device fitted to his car. When driving down the freeway it is set, say, at 100 kph. The device constantly monitors the revolutions of the tail-shaft. When it detects variation from the pre-chosen setting of 100 kph it makes the necessary adjustments by either opening or closing the throttle. The vehicle thereby maintains a speed of 100 kph plus or minus some small margin of error (Figure 11). This is

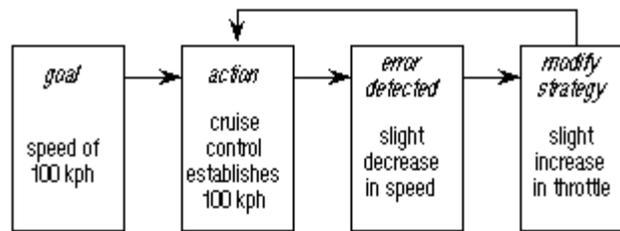single-loop learning. The speed limit functions as a governing value.

Fig. 11. Single-loop learning

Suppose we now approach a different speed zone where the maximum is 80 kph. The driver resets the control device to this speed and the car then maintains this new setting. The act of re-adjusting the speed from 100 kph to 80 kph is enabled by double-loop learning &endash; it is a higher order form of behaviour which seeks to modify the overall governing value (cruise control setting) in the light of changes in the overall environment.

In short, the cruise control is a system which displays single-loop learning. The goal (100 kph) is constant once set. The cruise control and the driver in conjunction form a double-loop system (Figure 8), because the driver may change the goal to adjust to the environmental conditions.

The relevance of the governing value is tested in double-loop learning. This is inimical to a Model 1 theory-in-use, as was illustrated earlier.
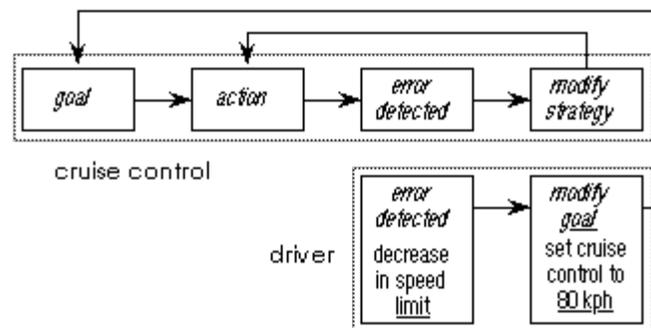

Fig. 12. Double-loop learning

Two other illustrations of the difference between single-loop and double-loop learning follow...

Imagine a couple in a relationship for quite some time. Like many such couples they have developed characteristic ways of dealing with problems which arise between them. Indeed, it seems as if this couple operate by a governing value which discourages the expression of emotionality and encourages rationality. Consequently, they believe it wrong to display anger or irritation openly and their disagreements are marked by extremely "polite" debate. Whenever it appears as though a disagreement is looming they take a cooling-off period and then come together to discuss it "rationally". This, again, is single-loop learning: the governing value of rationality is taken as a given.

Now imagine another couple in a similar situation who have used very similar strategies for resolving conflict. This time one partner says to the other in an uncharacteristic outburst: "Oh, if only you would get angry once in a while instead of having ice in your veins!". There follows a

heated exchange between the two, after which they sit down and talk about the value of expressing feelings as well as thoughts in conflicts, and the essentially-dissatisfying nature of their previous way of resolving differences. Thereafter, their disagreements are marked by both emotionality and rationality. This is double-loop learning: the underlying value of minimising emotionality has been examined openly between the two people.

Argyris suggests that Model 1 theories-in-use and the associated single-loop learning processes are acquired very early in life from our parents. In drawing on the work of Kelman (1958) he points out that most parenting processes reward compliance and identification. Parents reward children for behaving as the parents wish. What Argyris advocates as more effective is internalisation: the individual takes on a behaviour because it is intrinsically satisfying (Argyris and Schön, 1974, p. 82).

In doing this he is strikingly similar to Rulla et al (1978, p. 91), who also draw on Kelman's work. [6] They suggest that the intrinsic reward of internalisation [7] is obtained for the individual *by means of* her interaction with the environment. In interacting with the environment, the individual learns which behaviours achieve good outcomes. The behaviour is adopted because it works.

In other words, the interaction between the person and her environment "confirms" the validity of her underlying values. The person comes to experience a sense of satisfaction, generated by the match between her espoused governing value and the governing value of her theory-in-use.

On the other hand, according to Rulla, gratification through compliance or identification is gained *because of* interaction with the environment. The person does what is rewarded by others indirectly (we feel good when we imitate their behaviour) or directly (they reward us when we do what they want). The behaviour is adopted because it is similar to what others do, or because they reward it.

Argyris terms the beliefs about the environment the *field of constancy*. In single-loop learning this is treated as fixed because there is compliance or identification with the governing value. A dialectic between the actors and the field of constancy produces internalisation and, thereby, double-loop learning.

We have talked about similar issues in our own writing (for example Dalmau and Dick, 1991). Discussing the second watershed model, we make a distinction between traditional systems of socialisation and education, and some of the changed systems now sometimes in evidence.

Traditional schools and families and work settings teach people *solutions*: one-right-way methods of dealing with situations. Traditional etiquette is an example. The supporting beliefs are often held as beliefs without reason, adopted mainly through imitation.

Alternative systems, preparing for more rapid rates of change, teach *strategies* for reaching solutions. Individuals develop their own solutions by engaging with their environment. Encouraged to discover the appropriate behaviour, and the reasons for it, they are better able to adjust their behaviour in the light of changed situations.

Let us return to the case of the couple in the example. In the first example the governing values are reducing emotionality and maximising rationality. These constitute the field of constancy. They are Model 1, producing compliance and probably high levels of identification.

In the second example the unplanned heated exchange produces what Prigogine and Stengers (1985) would call a far-from-equilibrium position. This then triggers an exploration of the appropriateness of the governing values. From the resulting dialectic between the couple and their pre-existing values, they adopt a new equilibrium position, involving both new governing values and new strategies.

Many of the processes outlined in Part 3 of this book [not included here] are designed to help people and groups move to new equilibrium positions without having to go through what our second couple experienced, a far-from-equilibrium position (though, as we have argued, with some discomfort).

Now to our second illustration... Imagine an organisation in which there is evidence of increasing covert operational doubt (Dalmau and Dick, 1990): that is, members are beginning to identify increasing problems in the way, say, personnel policy is administered. There is dissatisfaction about late payments of the additional allowances people are owed, and over-legalistic interpretations of regulations. There is a general feeling that the Personnel Department is more of a hindrance than a help when it comes to providing due care and benefits to employees.

The Personnel Department, let us suppose, is organised by a manager who believes in the strict interpretation of rules and regulations, the zealous prevention or correction of minor fraud, and the resolution of all personnel disputes through formal correspondence and rational investigation. (We assure the reader that such units exist in many organisations, even today.)

This is clearly an organisation that works on Model 1 theories-in-use. Consequently, it is very difficult for employees to question openly the performance of the Personnel Department. If they do so, they violate some of the Model 1 governing values which maintain this very system: values which include, for example, be rational, and minimise expression of negative feelings.

We might expect a lot of grumbling in the corridors about the Department, but no explicit examination of its performance. Suggestions of poor performance may be made obliquely to a senior executive under whose wing the personnel function falls, with the inference that she should "fix it". There might well be a few polite suggestions to the manager of the Personnel Department to improve the unit's performance. But that may be all.

As Argyris (1982b, p. 7) points out, processes are often used which lead to immediate success and long-range problems. The complainants' feelings are acknowledged, perhaps. There is some minor adjustment of the personnel function. And before long the grumbling starts again. This is again single-loop learning.

Now imagine another organisation experiencing similar dissatisfaction, with quite different governing values. The senior executive is made aware of the grumbling. She involves all relevant stakeholders in immediate inquiry. This inquiry examines the personnel practices and procedures and *also* their overall purpose and direction (Dalmau and Dick, 1990). In other words, their governing values.

Valid information is sought about actual performance. There is public testing of this information and public questioning of the basic goals of the personnel function (Argyris, 1982b, p. 19). Threatening issues are surfaced and dealt with openly and honestly. There is a deliberate attempt to minimise camouflage of error. In effect, a dialectic is created; the open expression of dissent allows the identification of new goals and new responses (Argyris, 1982a, p. 106). Double-loop learning is achieved.

[index]

Achieving a learning system

Consider, now, how one might actually construct an organisation or other social system characterised by double-loop learning.

As noted above, systems with stable environments tend to use traditional socialisation processes,

with *identification* as one of the mechanisms. (This, you will recall, is when people behave in certain ways through imitation, or for external rewards.) Much of the literature on organisational culture which appeared in the early 1980s was written almost as if external stability were assumed. Its authors proposed building cohesive and united cultures based on "core values". It is instructive to consider the implications of these propositions for learning.

Strong cultures seem most likely to occur when identification is the primary method for generating collective identity and unity. We can call this an "identification culture".

Internalisation, however, poses some difficulty for the corporate social architect seeking to develop unity and identity. Each individual, as you will recall, generates her own self-satisfying governing values. If these overlap substantially then a corporate culture (an "internalisation culture") results. If not, we have either a series of sub-cultures or a collection of unconnected individuals.

As already mentioned, identification cultures tend to produce single-loop learning. Members may experience a strong sense of belonging to the group, but are less likely to recognise critical internal inconsistencies or external changes. [8] Their identity and long-term goals are likely to be taken as given, even immutable. Such cultures may continue until members suffer what Berger and Luckmann (1966) call a breakdown in plausibility. This collective phenomenon, which we have elsewhere [9] called "ethical doubt", corresponds to individual cognitive dissonance.
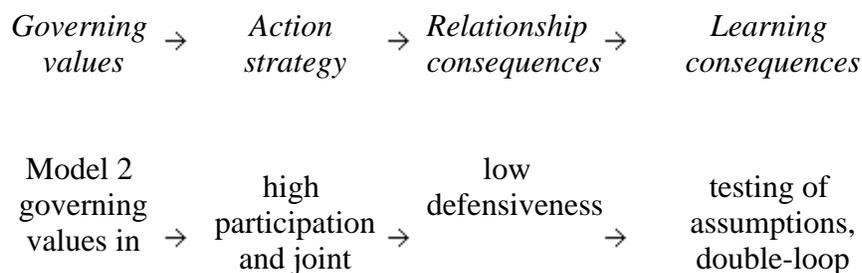
At the time of writing the first edition the Queensland National Party culture seemed to fit the description. For many years it had been highly cohesive. It succeeded in inducing extremely high levels of identification in members. It operated on "given" beliefs, seeking over a long period to make small single-loop adjustments to policy and practice in order to maintain the dominant culture. Then it began to fall apart. Events revealed discrepancies between espoused theory and theory-in-use. Eventually, the plausibility of both the party and the electorate was stretched beyond tolerance.

An internalisation culture, even if less cohesive, is more responsive to its external environment. It can transform dramatically over time as its underlying identity and unity undergo successive changes through double-loop learning.

How might the actual change come about? One possibility is to change communication styles. Argyris has shown, particularly in *Action science* (Argyris and others, 1985), how far reaching the use of open and non-defensive interpersonal styles can be in their wider effects. There is something to be said, though, for starting at the other end too &endash; creating *systems and structures* where openness, non-defensiveness and the like are easy, and encouraged.

We suspect Argyris would claim Model 1 patterns are too ingrained: changing the environment may not suffice. In any event, we believe so. But we have met people who *do* try to be open to change, only to find that the system makes this very difficult. Why not start at both ends?

Some of the features of a Model 2 theory-in-use might be summarised like this...

Governing values	→	*Action strategy*	→	*Relationship consequences*	→	*Learning consequences*
Model 2 governing values in	→	high participation and joint	→	low defensiveness	→	testing of assumptions, double-loop

| common use | responsibility | and high risk-taking | learning |

All (!) you have to do is create the Model 2 governing values. Moving towards an internalisation culture is one way. Controlling outcomes and resources rather than actions and procedures is how you might do so. There are other enticing possibilities too.

One is to change actions without controlling them. As it happens, attitudes and values have three components: beliefs, feelings, and action. If you change any one of these, you are likely to change the others. You can change the actions by first changing beliefs and feelings. You can change the beliefs and feelings by first changing actions. Either can create dissonance and thus motivate change.

On balance you will probably find it as easy to use the second of these; perhaps even easier. You change actions, by and large, by changing the rewards and penalties in the environment &endash; the internal rewards and penalties which feelings provide, or the external rewards and penalties delivered by other people, or both.

An environment which rewards participation, joint problem-solving and openness can be expected in the long term to move values more towards Model 2. (For this reason we have included an application on structuring social systems in Part 3 of the book from which this is extracted.)

For example we have been experimenting with the use of participative and qualitative evaluation models and their effects on performance. In particular we have been helping people to adopt short-cycle evaluation which requires a questioning, feedback-sensitive mindset. Our intention has been to make evaluation something that each person does for herself. It is not something left to the experts.

These evaluation models encourage people to give more attention to regular incremental improvement. [10] In general, the means which so far suggest themselves are...

- making provision for ongoing monitoring whenever plans are being developed or decisions made;

- setting up organisations or groups or programs as self-improving systems, by ensuring that the most relevant performance feedback is available, without threat, to those who can best use it to change their performance.

In general, this is more likely to work if decisions are participative, responsibility is widely shared throughout the system, there is high freedom of choice, and experimentation and change are encouraged.

This concludes the minimal conceptual foundations for what follows. Before proceeding to the tools and applications, let us try to provide a simple overview of the concepts.

[index]

In summary...

On the one hand, people claim to hold certain beliefs, and claim to observe these values in their behaviour. This is their espoused theory. On the other, there are beliefs which are implied by their actual behaviour: their theory-in-use. The two are very often discrepant, but the person is not aware of this. If people become aware, dissonance results; and this in turn may trigger change in their espoused theory or their theory-in-use.

Certain beliefs tend to cluster together. People who believe in the pursuit of winning and the avoidance of losing, for example, tend also to believe in being narrowly rational, and minimising emotionality. Argyris and Schön have identified two clusters of beliefs which they call Model 1 and Model 2. Model 1 might be characterised as adversarial, competitive, and narrowly rational. Model 2 is more consensual, more open to change, and provides more opportunity for choice.

The two models also have differing consequences for people, for systems, and for learning. In particular, Model 1 allows only single-loop learning: learning only within fixed limits. Certain beliefs are maintained as beliefs without reason. They are unchallenged and in fact unchallengeable, even though people often don't understand why. Model 2 allows double-loop learning in which the overall values are open to challenge.

There are some people who advocate Model 2 values. In other words, their espoused theory is Model 2. Their behaviour, however, frequently is more consistent with Model 1 values. This implies a Model 1 theory-in-use.

If such people could become aware of the mismatch, they might become motivated to do something about it. Unfortunately, however, this is difficult. The prevailing culture is Model 1, and so are many of the organisations and social systems within it. Further, anyone trying to inform them of the mismatch is likely to use Model 1 behaviour to do so.

In consequence...

- It is common for people to practise (through Model 1 theory-in-use) something other than what they preach to themselves (their Model 2 espoused theory).

- They are almost always blind to the mismatch. If not, they would do something about it.

- In a Model 1 culture, other people may notice the discrepancy, but are unlikely to bring it to their attention; to do so violates Model 1 taboos. And there are further taboos against acknowledging the taboos.

- If someone does attempt to bring a discrepancy to someone else's attention, it is likely to be done in a Model 1 way, and thus trigger a defensive interaction.

Relationships can often be characterised as defensive, such that people hold untested assumptions about each other. Relationships of whatever form are the building blocks of social structures. Structures influence the way in which people relate to each other.

It has been our wish to develop from this material some detailed processes for ease of learning and application. To do so, we have tried to specify the type of information typically available within Model 1 and Model 2 relationships.

To this end, we have described a sequence of events in which...

- actions of one person

- have material consequences for another person.

- In response, the second person develops beliefs, often in the form of assumptions about the first person's motives,

- and also feelings about the first person.

- Without checking out these assumptions, the second person intends to respond to the first person as if the assumptions were true

- and to some extent, responds as intended.

- This reaction is, in turn, an action which has consequences for the first person, who develops beliefs and feelings and intentions accordingly.

It is not unusual for a double self-fulfilling prophecy to be created in which each person's actions are taken as evidence for the incorrect assumptions which began it all. Diagrammatically...

$$\text{actions} \begin{matrix} \to \text{outcomes} \to \text{beliefs} \to \text{feelings} \to \text{intentions} \to \\ \leftarrow \text{intentions} \leftarrow \text{feelings} \leftarrow \text{beliefs} \leftarrow \text{outcome} \leftarrow \end{matrix} \text{actions}$$

The beliefs and feelings are seldom expressed. Often, neither are the outcomes or intentions. The people unwittingly deny each other information &endash; information which would enable them to develop a better understanding of self and other, and a more constructive relationship.

In general, people prefer to practise what they preach. When they become aware of a mismatch, they experience negative feelings. Unfortunately, they also experience negative feelings when they or another person violates the social rules of the social system and culture. Currently, most systems encourage Model 1 values.

To create change these conditions must be met...

- To isolate the people from the prevailing single-loop learning and taboos of the wider culture, a supportive climate must be created.

- To enable the dissonance-arousing mismatches between preaching and practice to become apparent, specific information about their behaviour must be provided.

- To generate this information, their informants must know which information is required.

- To escape the conventional taboos, provision of this information must be legitimised in some way.

- The process by which the information is given will work better if the informant does so in a Model 2 way.

- The process must also encourage the receiver to understand what is said.

In short, double-loop learning will more probably occur if both people engage in a mutual pursuit of honest information and understanding. In a sense, the processes can simulate Model 2 conditions, allowing the participants to discover Model 2 theories-in-use for themselves.

The applications which appear in Part 3 are intended to achieve this outcome. The tools in Part 2 are used as components of those applications or can be used in processes which you yourself design.

[index]

Notes

1. This is taken from the first chapter of Bob Dick and Tim Dalmau (1999) *Values in action: applying the ideas of Argyris and Schön*, second edition. Chapel Hill, Qld.: Interchange. [back]

2. In this respect their approach shows some parallels to the cognitive psychotherapies. [back]

3. Strictly speaking, the origins of intention are more likely to be the beliefs rather than the feelings; the arrows in the chain are therefore to be regarded as approximate representations of what goes with what. [back]

4. Grinder, Michael; Grinder, John; and Stephens, Laurie, eds. (1991), Righting the educational conveyor belt. Portland, Or.: Metamorphous. [back]

5. To be accurate, they use "Model I" and "Model II". [back]

6. There are also similarities here to the work of theorists on moral development. The seminal examples were provided by people such as Jean Piaget (1965), The moral judgment of the child, New York: Free Press. The earliest edition of this, we think, was 1932. In the US much of the work is underpinned by that of L. Kohlberg (1969), Stage and sequence: the cognitive-developmental approach to socialisation, in D. Goslin (ed.), Handbook of socialisation theory and research, Chicago: Rand McNally. [back]

7. In Rulla et. al. pp. 222-223, identification is defined as "the process of accepting influence from a social agent or referent, based on the self-defining value of the relationship with the social agent, person or group." Internalisation is defined as "the process of adopting a way of behaving or thinking because it is congruent with one's value system." [back]

8. You will recognise this as what Irving Janis has called "groupthink". See I. Janis (1972), Victims of groupthink: a psychological study of foreign policy decisions and fiascoes, Boston: Houghton-Mifflin. [back]

9. For example Dalmau and Dick (1991), p. 9. [back]

10. The notion of continuous incremental improvement is the heart of the Japanese approach to quality management, currently being adopted in the West because of its obvious success in Japan. See for example Masaaki Imai (1986), Kaizen. Short-cycle evaluation, we think, is a more Model 2 alternative. It doesn't depend so heavily on standardised procedures. [back]

References

Argyle, M., Furnham, A. and Graham, J.A. (1981), *Social situations*. Cambridge University Press.

Argyris, C. (1976), *Increasing leadership effectiveness*. New York: Wiley.

Argyris, C. (1982a), *Reasoning, learning and action: individual and organisational*. San Francisco: Jossey-Bass.

Argyris, C. (1982b), The executive mind and double-loop learning. *Organisational dynamics*.

American Management Association, Amacom Periodicals Division.

Argyris, C. (1985), *Strategy, change and defensive routines*. Marshfield, Mass.: Pitman.

Argyris, C. and Schön, D. (1974), *Theory in practice: increasing professional effectiveness*. San Francisco: Jossey-Bass.

Argyris, C., Putnam, R. and Smith, D.McL. (1985), *Action science: concepts, methods and skills for research and intervention*. San Francisco, Ca.: Jossey-Bass.

Argyris, C. and Schön, D. (1978), *Organizational learning: a theory of action perspective*. New York: McGraw-Hill.

Ashby, W. (1952), *Design for a brain*. New York: Wiley.

Berger, P. and Luckmann, T. (1966), *The social construction of reality: a treatise in the sociology of knowledge*. N.Y.: Doubleday.

Dalmau, T. and Dick, B. (1990), *A diagnostic model for selecting interventions for community and organisational change*, revised edition. Chapel Hill: Interchange.

Dalmau, T. and Dick, B. (1991), *Politics, conflict and culture: a journey into complexity*, revised. Chapel Hill, Qld.: Interchange.

Dick, B. (1986), *Learning to communicate*. St Lucia: University of Queensland Bookshop, and Interchange.

Egan, G. (1983), *Change agent skills, I* (preliminary manuscript). Chicago: Brooks-Cole.

Fishbein, M. and Ajzen, I. (1975), *Belief, attitude, intention and behaviour: an introduction to theory and research*. Reading, Mass.: Addison-Wesley.

Grinder, Michael; Grinder, John; and Stephens, Laurie, eds. (1991), *Righting the educational conveyor belt*. Portland, Or.: Metamorphous.

Imai, M. (1986), *Kaizen: the key to Japan's competitive success*. New York: Random House

Janis, I. (1972), *Victims of groupthink: a psychological study of foreign policy decisions and fiascoes*. Boston: Houghton-Mifflin.

Kelman, H. (1958), Compliance, identification and internalisation: three processes of attitude change. *Journal of Conflict Resolution*, 2, 51-60.

Kohlberg, L. (1969), Stage and sequence: the cognitive-developmental approach to socialisation, in D. Goslin (ed.), *Handbook of socialisation theory and research*. Chicago: Rand McNally.

Miller, S., Nunnally, E.W. and Wackman, D.B. (1979), *Talking together*. Minneapolis, Mann: Interpersonal Communication Programs.

Piaget, J. (1965), *The moral judgment of the child*. New York: Free Press.

Prigogine, I. and Stengers, I. (1985), *Order out of chaos*. London: Flamingo.

Rulla, L., Imoda, F., and Ridick, J. (1978), *Psychological structure and vocation*. Dublin: Villa Books.

Rulla, L., Ridick, J. and Imoda, F. (1989), *Anthropology of Christian vocation: Volume II: Existential*

confirmation. Rome: Gregorian University Press.

———

This document can be cited as follows:

Dick, B. and Dalmau, T. (2000) *Argyris and Schön: some elements of their models.* [On line]. Available at
http://www.uq.net.au/action_research/as/argyris2.html

Maintained by Bob Dick; this version 1.01w last revised 20000106