EVALUATION AND SYSTEMS THINKING

Read Me First

Health Warning – January 2006

The core of the document was written in early 2002. A lot has happened to my understanding of systems theories and practices since then. Much of it has yet to find its way into this document. Other documents on my web site explore individual systems approaches and reflect much more the current state of my knowledge. From time to time I make minor adjustments to this document, but the hoped for Version 2 mentioned below is not likely. However, your feedback is still welcome and will find its way to other things that will be just as publicly available.

What this draft is all about

This is an early attempt to introduce evaluators to the diverse work of systems based approaches to inquiry and development. The draft seeks to be accessible and highlight the core principles. What this draft will become is still uncertain, and will emerge. I'd like it to become a core for a workshop, and I'd welcome ideas about how to do this. On the other hand, you may have other suggestions for other formats.

Essentially this draft contains a selection of systems based theories, identifies some strengths & weaknesses of these theories, and explores potential implications for evaluation. In this draft I've made no major attempt to draw conclusions. Also, given the complexity of both "evaluation" and "systems" as fields of inquiry, I've made no attempt to be either comprehensive or complete. I have however, tried to cover the range from the mechanical to the (almost) metaphysical. I've also tried to restrict discussion on each systems approach to four or pages at most. Hopefully what gets lost is compensated by the questions that such brevity invites.

The descriptions are often very crude, and nearly always second hand. The commentaries on strengths and weaknesses, and implications for evaluation are very tentative. I would like them to be more insightful, and hope that is what develops as people read this draft.

What I'd like you to do

This is intended to be a developing document. To be read, commented on and changed. I'd like it to become an "our" document, rather than a "my" document.

However, I'd like to put some boundaries around this. We can debate these boundaries of course, but these are the boundaries that informed this current draft.

- This work is ultimately intended for an audience of relative beginners who wish to scan the field and make a choice about their own approach to systems thinking
- I don't want this draft to become a focus for a round of methodology, method or technique wars. Some people who read this will know a lot about one approach and feel passionate about it. That's fine – I've my favourite approaches too. However, I was involved in a previous attempt like this which almost fell to bits because respondents often wanted someone else to write *their* paper on *their* subject, from *their* viewpoint for *them*. Rather than focussing on convincing ourselves and others about a particular approach, I our job as helping to enlarge the evaluation field's overall understanding of systems thinking. If necessary we can fight over the spoils once we have 'em. In particular I want to undermine many evaluators' notion that systems approaches are little more than fancy versions of program logic or concept mapping.
  So …
- … I'd ask we try to keep a spirit of critical reflection within a dialogue framework.
- I'm happy to co-ordinate this endeavour with three conditions
    a. In the first instance, I really like to keep the three part structure (see below). I'd therefore appreciate initial comments within that structure. That will help me get over the first hump. After that, I'm open to suggestions.
    b. I want this to be collaborative. Sending me wodges of material to read and assimilate is not, in my view, collaborative. It's passing the buck back to me. [Not that I don't want you to send the stuff, just don't assume I'm going to do the hack work]
    c. I'm self-employed, and like you I have a living to make and a life to lead. I'm on the road February to May 2002 and I have no access to academic libraries or decent bookstores. Therefore, I'd prefer material and ideas that can be easily slotted in to the emerging structure of the endeavour.

Other than that let's have some fun with this, and see what emerges.

Finally let me know by email that you have downloaded this draft, and which version. I can then tell you about any subsequent drafts. Let me know if you pass it on to someone else.

Structure of this paper

The "systems" approaches covered are :-

- System Dynamics
- Soft Systems Methodology (SSM)
- Cultural-Historical Activity Theory (CHAT)
- Complexity theory
- Fifth Discipline
- Critical Systems Thinking
- Systemic Thinking

I've excluded several others (eg cybernetics, open systems theory, integrated planning) because I know less about them, rather than their particular merits. I've also excluded Program Logic and Program Theory, partly because of their familiarity to evaluators, but mostly because I don't think they constitute a distinct branch of systems thinking; they are primarily applications of particular systems approaches. But I'm persuadable.

The structure of each section is the same.

1. A description of the systems approach
2. Acknowledged strengths and weaknesses
3. The implications for evaluation

Two key definitions

One of the main problems is that both "systems" and "evaluation" have meanings and uses beyond the technical. The word "system" is commonly used interchangeably with "process". "Evaluation" can mean any form of judgement.

With no attempt to be universal, here are the meanings applied in this document.

Evaluation

Seeking to answer accurately, validly and usefully the following three questions :-

What happened (or is happening)?
So what ?
Now what ?

Or more precisely :-

"Who, in what circumstances, and in what way, did or didn't benefit from the thing you are evaluating, who learned what from that, and what does that imply for the future" ?

System

The origins of systems theory and many of the approaches described here draw from two interconnected threads. There were the "engineers" such as Von Bertallanfy and Ackoff, and there were the group dynamacists and organisational developers such as Bohm, Emery, Trist, Revans and to some extent Lewin.

Ackoff's classic definition has 31 properties of a system[1]. For simplicity however, I've used a short version developed by my old colleagues at the Open University in the UK[2] :-

A system is an assembly of parts where :-
1. the parts or components are connected together in an organised way
2. the parts or components are affected by being in the system and are changed by leaving it

3. the assembly does something
4. the assembly has been identified by someone as being of special interest

An awful lot of water has gone under the bridge since that definition was formulated. Its main disadvantage is that it encourages the idea that a system is a "thing", whereas many theorists influenced by later constructivist ideas (eg Checkland, Churchman, even Engestrom) have developed theories which rely on a "system" as something essentially conceptual. Indeed Checkland's definition of a system is, by his own admission and insistence, completely implausible. These approaches allow you to delve deeper into the contradictions between theory and reality in order to learn more about the real world. One the other hand, as we shall see, it can frustrate the Hell out of strongly sensate and normative managers and researchers.

SYSTEM DYNAMICS
[For a more detailed example, see http://users.actrix.co.nz/bobwill/AESSD.pdf]

What is System Dynamics ?

It has been remarkably difficult to find a simple description of what system dynamics actually is. Most descriptions say something about feedback loops and computer simulations and then launch into something terribly technical.

System dynamics was developed by Jay Forrester at MIT back in the 1960's from Stafford Beer's concepts of cybernetics. Its popularity has waxed and waned in ever since. It got a boost in the 1990's when Peter Senge took system dynamics, placed it alongside organisational and personal development concepts and called the result "Systems Thinking"

Essentially system dynamics is based on two notions about cause and effect. Firstly, in any system there is *detail complexity* – large number of potential causes and effects. Secondly and more critically, there is *dynamic complexity*. Not only does cause contribute to effect, but the effect in some way contributes to the cause. Furthermore, most systems have complex arrays of such feedback loops. What is important here, is that any delay in a process or feedback loop when placed alongside all the others mean that the dynamic complexity increases to such an extent that most people will misinterpret what actually is going on.

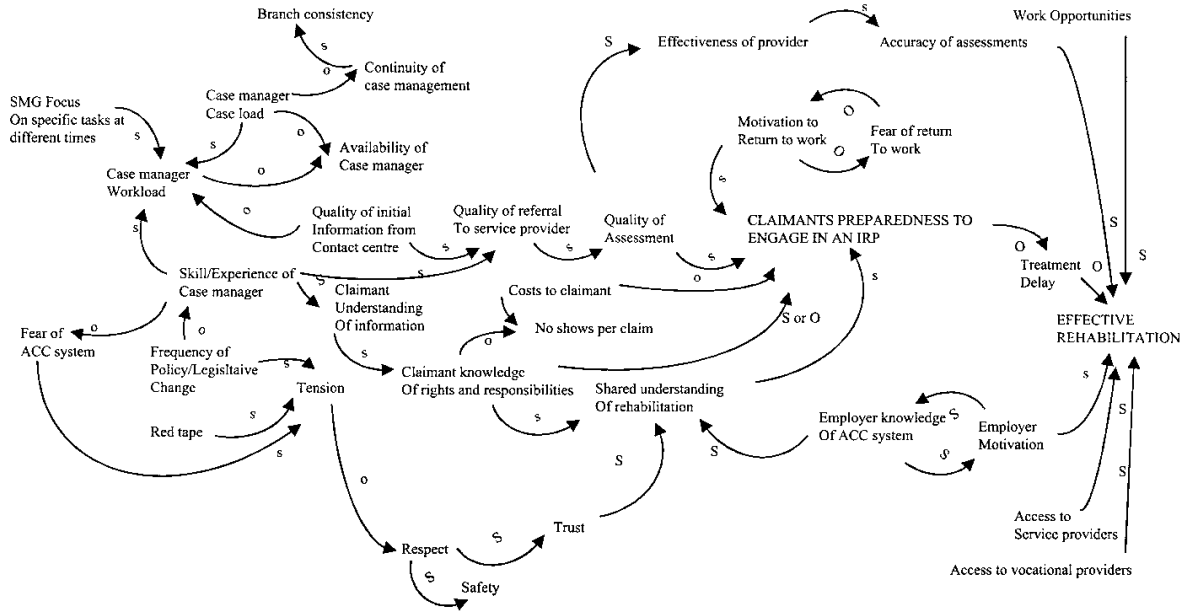
There is a famous simulation exercise of a brewery (the Beer Game) that shows the greater importance of dynamic complexity over detail complexity. In this simulation, there is very little detail complexity – it is a simple demand and supply chain (brewing beer for customers). But dynamic complexity is added by the delay caused by having to wait two weeks for the beer to brew. When asked to "manage" this system, even skilled managers generally perform worse than if they took their decisions by throwing a dice. And this is a really simple system – much less complex than what most managers actually seek to manage

So System Dynamics focuses primarily on identifying the main variables of a system (ie the bits that are able to change) and exploring rigorously the effect they have on each other. It therefore emphasises the "dynamics" of the system, rather than seeking a snapshot of the system.

There are many approaches to developing a system dynamics model but invariably they all end up with a mess like the following diagram of an injury rehabilitation system*.

[For the purposes of this exercise, don't try and make sense of the diagram]

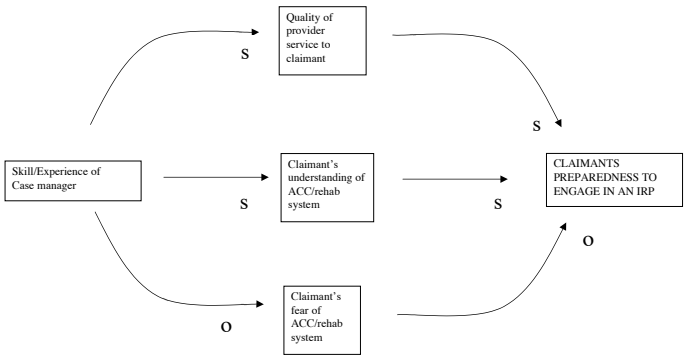
* Actually this an interesting example, since there are very few feedback loops. This was the point we made to the clients – this particular system was almost completely dependent on things outside the system. In other words it was, from the perspective of someone "inside" the system, almost completely out of control.



At this point you can do two things (apart from breaking out the paracetamol).

1. You can take all the variables and their inter-relationships, plug them into a computer simulation program and play lots of "what ifs".

2. You can seek to stand back and try to understand the essence of what is going on. This doesn't have to be in the form of a diagram, but it often is. Here's an example from the above diagram. We simplified the diagram so that it demonstrated the importance of synchronising three parallel threads of the injury management process; one administrative, one emotional and one cognitive.



System archetypes

Over the years, some generic "simplifications" have been developed with catchy names like "fixes that fail", "success to the successful", "the tragedy of the commons", "shifting the burden", "eroding goals".

Peter Senge in his book "5th Discipline"[3] popularised these "system archetypes". Skilled system dynamics people can spot these dynamics in quite complex systems. In the above case, the client had been trying to fix the problem by focussing primarily on one thread – an example of a "fixes that fail" archetype.

Strengths and Weaknesses of System Dynamics

Strengths

- It is specifically designed to look at the dynamics of a system quantitatively, even though it may be based on a qualitative analysis
- You can test and develop hypotheses
- The debates around the relationships between the systems components usually exposes deeply held but often untested assumptions about the way the system works.

Weaknesses

- If you don't go the simulation route (ie you map the system but don't model it) you can't explore how all the variables interact with each other dynamically. For instance the "simplified" example shown above could be wildly wrong. Not only are we unsure that we have identified the most important variables, we don't for certain know the actually scale of their relative dynamics within the system. The Beer Game for instance looks (ie maps) really simple. People usually feel very confident in predicting what to do to get a stable brewing process. Yet when you simulate it (ie model it) most people take the wrong decisions and the system fluctuates wildly.
- Some people get obsessed with the need to include absolutely every possible variable; essentially confusing simulation of reality with modelling to explore and expose assumptions.
- In deeply qualitative situations, the task of getting accurate and valid data for modelling can require a great deal of expertise. Often entirely new information systems need to be set up to collect the data.
- According to Bob Flood, it doesn't model emergent phenomena very well, although it may well model the emergent process[4]
- It doesn't grapple very well with the problem of what to do when the whole thing expands beyond the realm of investigation. There are no rules for boundary setting.
- It isn't inherently "reflective", although most people would use some form of reflective process in assessing what the model says.

Implications for evaluation of System Dynamics

Most evaluations, at least in my experience are essentially descriptive. At best they map a system and its environment, (eg program logic, concept mapping), rarely do they try to model it. Yet "now what ?" is a critical part of evaluative activity –

evaluators pose themselves a modelling question yet often base the answers on a mapping analysis of the system. System dynamic approaches could provide some opportunities here.

Most clients of evaluations (and many of the stakeholders) seek certainty. Often they seek confirmation of their world view. System dynamics often increases uncertainty, and often exposes so-called information systems to be woefully inadequate for the task they are supposed to do. All evaluations can be threatening, but system dynamics approaches can be more threatening than most. The client on the receiving end of the diagrams shown earlier didn't know what to do with them, and I suspect didn't quite trust the process enough to consider the insights that flowed from it. Using system dynamics requires re-educating clients.

It's a highly technical process if done to the simulation stage. Few evaluators have these skills – there is a skill issue.

SOFT SYSTEM METHODOLOGY

For a longer description refer to the following part my of my WEB page
http://users.actrix.co.nz/bobwill/ssm.pdf

What is SSM ?[5]

Soft Systems Methodology (SSM) was developed by Peter Checkland in the late 60's at the University of Lancaster in the UK. Originally it was seen as a modelling tool, but in later years it has been seen increasingly as a learning and meaning development tool.[6]

At the heart of SSM is a comparison between the world as it is, and some models of the world as it might be. Out of this comparison arise a better understanding of the world ("research"), and some ideas for improvement ("action").

In SSM the researchers begin with a real-world problem. They study the systems which contain the problem. Following this, they develop some models of that system that generate insights into the way the "real-world" problem might be addressed. As SSM is a systems methodology, the models *must* be formed using systems rules and concepts[7].

These models are then compared to the actual situation. Differences between the models and reality become the basis for planning changes.

The "classic" methodology has seven stages

1 The problem situation unstructured

The problem situation is first experienced, as it is, by the researcher. That is, the researcher makes as few presumptions about the nature of the situation as possible.

2 The problem situation expressed

In this step the researcher develops a detailed description, a "rich picture", of the situation within which the problem occurs. This is most often done diagrammatically.

In addition to the logic of the situation, the rich picture also tries to capture the relationships, the value judgments people make, and the "feel" of the situation.

3 Root definitions of relevant systems

Now the "root definitions", the essence of the relevant systems, are defined.

There are two aspects, the logical analysis and the cultural analysis.

For the logical analysis, Checkland provides the mnemonic CATWOE as a checklist for ensuring that the important features of the root definitions are included:

Customers.....who are system beneficiaries
Actors.....who transform inputs to outputs
Transformation.....from inputs into outputs
Weltanschauung.....the relevant world views
Owner.....the persons with power of veto over the system
Environmental constraints...that need to be considered

The "transformation" element is one of the features that signal this as a "systems" approach.

Checkland encourages you, indeed virtually insists, that you develop several "root definitions". Each "root definition" will have different elements in CATWOE. You are expected to be creative. For instance think of a professional soccer match. What root definitions flow from seeing it as a means of keeping hooligans out of shopping malls, police training for crowd control, a means of keeping 22 men fit and well paid, a way of entertaining 100,000 in one place, an exercise in product advertising ? A root definition could be "A system that permits police recruits to be more highly trained in the humane control of a large excited crowd, allowing them to keeping within current legislation and out of the Sunday newspapers and the Office of the Police Superintendent." Here the Customers are the police recruits, the actors are the crowd, the transformation is the ability to control crowds, the Weltanschauung is "humane" (in evaluation terms the value or worth), the owner is the Superintendent, and the environmental constraints are legislation and the Press. Not all "root definitions" have to be this cumbersome.

The cultural analysis has three parts:

- A role analysis, focusing on the intervention itself. This seeks to identify the client, the would-be problem solver (the researcher), and the problem owner (roughly, stakeholders). In the terms that we used in earlier sessions you could think of this as the diagnostic part of entry and contracting.

- A social system analysis. This identifies, for the problem situation, three sets of elements: roles, norms, and values.

- A political system analysis. This identifies the use of power in the problem situation.

4 Making and testing conceptual models

This is a critical and rigorous step. The task now is to develop a systems model using *only* those elements of the root definition, in a way that *flows logically from* that root definition, and has all the properties which define a system (Checkland has his own definitions of key system components[8]). Other systems theories may also be used to test and reframe the conceptual model.

The researcher now draws upon their knowledge of systems concepts and models. They develop descriptions, in system terms, of how the relevant parts of the situation might function.

5 Comparing conceptual models with reality

The purpose is not to implement the conceptual models. Rather, it is so that models and reality can be compared and contrasted. The differences can be used as the basis for a discussion: how the relevant systems work, how they might work, and what the implication of that might be.

6 Identify feasible and desirable changes

From the discussion at step 5, certain possible changes are identified. They are likely to vary in desirability and feasibility:

- desirable: is it technically an improvement?
- feasible: especially, does it fit the culture?

7 Action to improve the problem situation

The most desirable and feasible changes identified at step 6 are now put into practice.

In the past decade, Checkland has modified this basic structure into four core activities[9] :-

a. Finding out about a problem situation
b. Formulating some relevant purposeful activity
c. Debating the situation using the models; seeking from the debate desirable and feasible changes which would improve the situation, and accommodate conflicting interests
d. Taking action to improve the situation.

Coincidentally, Bob Dick has constructed an alternative version, that overcomes some of these problems. He reframes the 7 steps into 4 discussions :-

a. between immersion (the rich picture) and essence (the root definitions)
b. between the essence (the root definitions) and the idealised (the conceptual models)
c. between idealised (conceptual models) and reality
d. between plans and implementation[10]

Strengths and weaknesses of SSM

Strengths

- It is potentially highly creative, but within a rigorous framework
- It actively promotes alternative views of exploring the "purpose" of the system
- When Step 5 is done well, it is powerfully reflective.

Weaknesses

- The mechanics of the process can overwhelm the spirit of the process
  In its traditional form, it is very demanding intellectually. Many people find it impossible to make the leaps from Step 2, which is grounded in data, to Step 3 which is essentially conceptual, and then back to reality in Step 6.
- It is very easy when comparing the model with reality, to assume that the model is the "ideal" system rather than an "idealised" one. Rather than gain insights in the comparison, you seek to make "reality" into the "idealised". As Checkland says, this is impossible.
- Earlier versions of the model don't really address issues of power and knowledge distribution. Later versions include the cultural analysis in stage 3, that become critical in stage 6 when determining the feasibility and desirability of the proposed changes to "reality".

Implications for evaluation of SSM

Frankly I'm curious why SSM isn't commonplace in evaluation. It falls well within most evaluation briefs, and evaluation definitions. CATWOE covers virtually everything that an evaluator needs. When you rename weltanschauung as values (to get the mnemonic VOCATE) the process becomes strongly value centred. It is stakeholder oriented (CAO). It is process and change oriented (T). And it considers environmental constraints (E). It has lots of rigour, and you are constantly testing the validity of your data and conclusions. It is strongly learning focussed (if done well) and focussed on system improvement.

So why haven't even I used it very much ? There are probably several reasons :-

1. It is intellectually very demanding. To get a robust and valid conceptual model (rather than a wiring diagram) you need to know your systems rules very well.
2. Parts of the process are highly conceptual, yet demand lots of discussion with stakeholders. Unfortunately many stakeholders (and many evaluators) are practical and pragmatic and want models to be more "realistic".
3. It's risky. I tried to use it in my PhD and after three years the whole endeavour fell apart – mainly because of reasons 1 and 2.

On the other hand, many practitioners who use Program Logic or Program theory based approaches will recognise some of the steps, if not the systems components and their rigour. In fact, I've occasionally used Program Logic or Concept Mapping type approaches as a less rigorous surrogate for Steps 3 and 4, and proceeded to take these into SSM stages 5 and beyond.

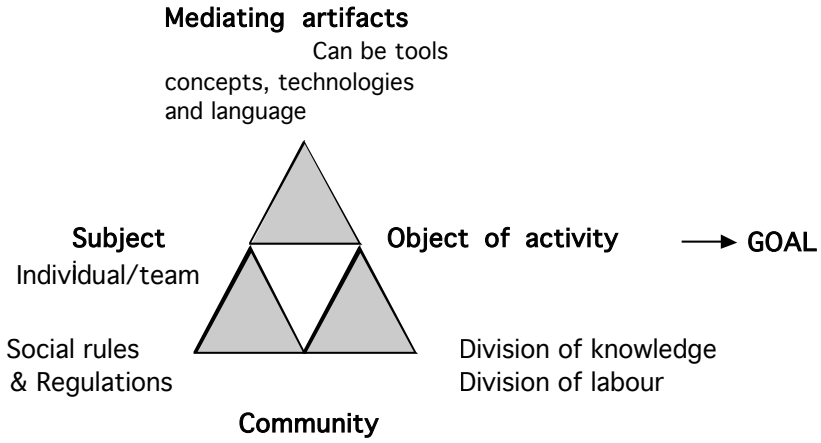
CULTURAL-HISTORICAL ACTIVITY THEORY (CHAT)

For a longer description refer to the following part my of my WEB page :-
http://users.actrix.co.nz/bobwill/activity.pdf
I'm also indebted to my colleagues at WEB Research -
http://www.webresearch.co.nz - whose expertise I've drawn on.

What is CHAT ?

CHAT was originally developed by Yrjo Engestrom at the University of Helsinki. According to its proponents, the theory essentially covers three things. Whilst each component is quite well known, the important aspect of this theory is that it seeks to accommodate the relationship between all of them.

1. It describes the factors that affect or mediate between individual behaviour (the "subject") and the results of that behaviour (the object) in order to achieve the "goal" in the context of wider social systems. These wider factors are not merely "external context", since they are part of the system - changing as a result of the system's activities.



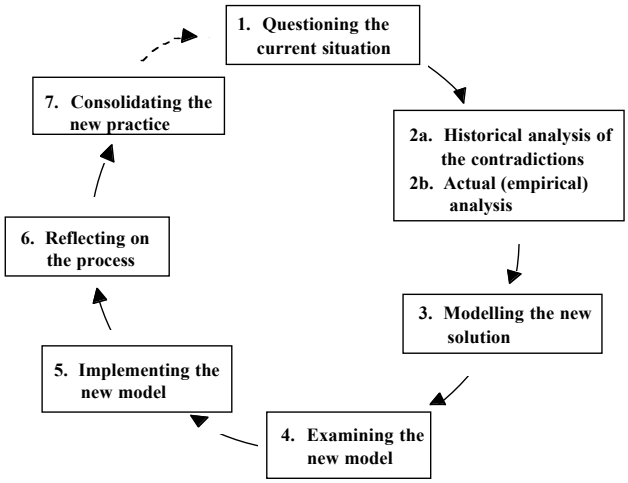
There are three key relationships, and three mediating factors. The clear triangle shows the key relationship between the individual, the community of which they are part and the activity they are all engaged in. The three shaded triangles represent the relationships that mediate the relationship between individuals, activities and communities. So :-

- concepts and technologies mediate (ie they have an active influence upon) the relationships between the individual and the object of his or her action, and vice versa;
- social rules and regulations mediate the relationships between the individual and the community, and vice versa;
- and a division of knowledge and labour mediate the relationships between community members and the shared object of the activity system, and vice versa.

2. CHAT proposes that these features can be aggregated (ie the "individual" can be a single person, a group of people, an organization or a community). So for instance, the "activity system" that is a branch office of a government agency,

influences and is influenced by the wider activity system that is New Zealand society.

3. It contains a theory of learning that explains how an individual (or groups of individuals) interact with these factors *so that they learn from that experience*. The theory of learning is based on Vygotskyian concepts of identifying and explaining "patterns" and "puzzles or contradictions". The diagram below has some similarities with Kolb's learning cycle, but, it is argued, is more directed and expansive.



Source: Expansive Learning Cycle; Engestrom 1996

Strengths and Weaknesses of CHAT

Strengths

- It's very close to a theory of everything !
- Unlike many systems models which emphasise "learning" as critical component, CHAT is actually driven by a particular learning theory.

Weaknesses

- Superficially, and in its parts, it seems to state the obvious. It's not until you really explore the implications of the model, that you begin to understand its collective power.
- It's not obvious how to translate the model into practice, although many studies have been done using the model. As yet there isn't a simple how-to-do-it-in-practice guide that is widely available.

Implications for evaluation of CHAT

At one level much evaluation practice unconsciously parallels some aspects of CHAT. Evaluators do explore the components of CHAT quite explicitly, as well as the way in which they interact with each other. Many seek explanations for why things happen that follow the two main frameworks quite closely.

There are two areas worth considering more deeply.

Firstly, there are parallels between CHAT and Realistic Evaluation methodology. Both seek to explore and explain generalisations and exceptions within a context. Both are strongly oriented towards creating and testing hypothesis. What CHAT does though is add both a degree of discipline *and* a theory of change. Where Realistic Evaluation seeks to explain via mapping, CHAT seeks to explain via a form of modelling.

The second potential contribution is in the area of evaluation capacity building and utilisation. If you think of evaluation as a "tool" within a CHAT construct, then CHAT provides both an explanation of why "the system" reacted to an evaluation the way that it did, and also a model of how best to get people to reflect deeply and learn (ie the cycle of expansive learning) by exploring and explaining collaboratively patterns and puzzles.

COMPLEX ADAPTIVE SYSTEMS THEORY (CAS)
For another example see the document on my website :
http://users.actrix.co.nz/bobwill/CASmaterial.pdf

What is CAS ?

CAS is one of a suite of theories that surround the concept of "complexity". The complexity field has developed an arcane set of concepts and jargon. However at the core there are some fairly simple concepts, that many claim (with some justification in my view) fundamentally alter the way we see the world and should react to it.

One way to explore CAS is to compare the features of so-called "simple" systems, "complex" systems, and their nemesis "random" or "chaotic" systems. Eoyang and Goldstein have come up with this table[11] :-

	SIMPLE	COMPLEX	RANDOM
Metaphor	Machine	Living Organism	Ricocheting Bullets
Leadership Roles	Predict & Control	Watch for :-Patterns Differences Connections Innovation	Fight Fires
Examples	Factory	High Performing Teams	Phone calls
Relationships types	Directive	Empowering	Disinterested
Structure	Hierarchy	Distributed Network	No Pattern
Future/Outcome	Predictable	Both Predictable & Unpredictable	Not Predictable
Boundaries	Rigid	Fuzzy & Changing	None
Feelings of people within these systems	Security & Boredom	Anxiety & Excitement	Confusion & Fear
Conflict is generally ….	Destructive	Constructive	Unproductive
Problem Focus	Parts	Whole	Shifting Collection
Decision are taken on the basis of ….	Facts, Data	Emerging Patterns	Instinct
Useful When	Routine Situations	New Opportunity	Unrelated Issues
Systemic Change created via …	Pressure	Patterns & Instability	Chance
Change Flow	Continuous/incremental	Precipitous/sudden	Erratic
Change Strategy	Push ~ Pull	Adapt & Self-organize	Hit ~ Miss
Response to unexpected events	Punish	Learn from	Cancel Out
Advantages	Efficient & Reliable	Adapt to Environment	Totally Free
Risks	Stagnation	False Starts	Irrelevance
Action	Manipulate	Participate & Relate	Be Acted Upon
Motivation	Compensation	Satisfaction	Safety

Fundamental Properties of Complex Adaptive Systems

But what creates these features ? What is going on underneath, deep in the complex adaptive system ?

One way of looking at CAS is through seven fundamental properties[12] :-

Butterfly Effect
Complex adaptive systems have extreme sensitivity to initial conditions. In other words a very small initial intervention can have a tremendous ultimate effect.

Boundaries
The area lying between different parts of the system are critically important. These boundaries don't have to be physical, they can be differences of viewpoint, experience, expectation, or culture.

Transforming feedback loops
Information that flows across these boundaries (if it can), has a significant and transforming effect compared with "simple" systems.

Fractals
These are simple rules, practices, procedures, principles or values that are present in every part of the system – no matter how big or small that part of the system is. Glenda Eoyang reckons that every CAS has no more than four or five fractals – at least one relating to something that makes a big difference, at least one relating to feedback, and at least one about boundaries.

Attractors
A complex adaptive system is made up of a large number of interdependent sub-systems. The interactions between them (ie the effect of the transforming feedback loops) create recognisable patterns. However, these patterns can only be understood at a large system level. They cannot be used to predict the behaviour of smaller parts of the system (ie the whole cannot be understood by measuring the performance of the parts).

Self-Organization
When pushed to an extreme state, complex adaptive systems without an obvious plan, and without the control of any single individual component spontaneously reorganises themselves. This is a controversial aspect of CAS, since it appears to challenge the current dominant organisational paradigm of charismatic leadership, strong management systems and explicit strategy. It means we are at best influencers of the system than managers of it (see below).

Coupling
Essentially this means that the relationship of the system and its environment, between parts of the system, and between individuals within the system is based on notions of "influence" rather than "control".

Strengths and Weaknesses of CAS

Strengths

- Complexity based theories seeks to be a "theory of everything"
- It has enormous potential explanatory power

Weaknesses

- Whilst complexity theory does seek to be a "theory of everything", many argue that it's some way off that yet.
- It strongly challenges current organisational practice and thinking, especially the particular contribution of leadership, strategy, long range planning and most traditional management processes (what some people call the "Boston School" of management theories).
- As yet there are few examples of using complexity theory in the kinds of areas that evaluators specialise
- The jargon that has grown up around complexity theory is highly technical and difficult to access.
- The tools by which to investigate complex systems are still being developed and tested
- By its very nature it doesn't offer the certainty and security that many evaluation clients are seeking. Indeed it challenges some of the fundamental beliefs about how "change" occurs.

Implications for evaluation of CAS

As yet few influential evaluators have explored complexity theory. It is little understood and not a little threatening to the strong input/process/result methodologies that dominate evaluation. However when confronted with the features of complex adaptive systems in the above table, many evaluators acknowledge that most of their work focuses around complex adaptive, rather than simple or chaotic systems.

According to Eoyang and Burkas[13] the implications for evaluators can be quite profound. Evaluations using CAS need to be dynamic. Because a CAS is dynamic, evaluation systems should incorporate flexible and dynamic features. Specifically, they should capture an emerging model of causal relationships, evaluate and revise the evaluation design often, capture, preserve and learn from the "noise" in the system. They need to be able to cope with the massive entanglement of the systems they are studying (eg incorporate multiple strategies, cycle times, time horizons, dimensions and informants). Evaluations need to be scale independent. Because a CAS incorporates many self-similar levels of organization, an evaluation program must incorporate both micro- and macro-patterns and structures. They need to be able to observe evolutionary change in individual and system-wide behaviour over the course of an assessment period. And they need to be able to respond to and observe emergent patterns.

As a consequence of this the focus, tools and techniques of evaluation will have to move from the structured, low-dimension, predictable patterns of much of traditional research to more organic and flexible strategies. They also need to provide more structure and pre-designed rationality than many of the individualistic and constructivist methods of qualitative evaluation. By including a wide range of approaches, CAS methods of evaluation can integrate the best of many disciplines and methods that were previously irreconcilable.

Many evaluators seek simplicity without being simplistic. Complexity theory in practice holds out that promise of discovering and determining what the few simple drivers of highly complex systems actually are. To some extent, that is a potential holy grail to evaluators drowning in detail and the specific.

But there is a skill issue. Complexity clearly poses some challenge to the tools evaluation uses. Evaluation is overwhelmingly dominated by qualitative and quantitative data collection and analytical methods developed within a "simple" system paradigm. There are exceptions (eg time series analysis, causal loops), but new skills will need to be developed.

There is an evaluator issue. Evaluators need to be able to absorb uncertainty and create learning environments for participants.

And there is a client relationship issue. When it comes to organisational and change theories, evaluation is a follower, not a leader. By its very nature it tends work within rather than challenge dominant investigative paradigm. Perhaps evaluation must wait until its primary clients (ie those designing, developing and running products and programs) have begun to embrace complexity.

"5th DISCIPLINE"

I thought long and hard about including this approach to systems. In some ways it is like Program Theory or Program Logic – an application of systems theory rather than a particular systems theory. Indeed, its originator – MIT's Peter Senge – has never claimed it to be a systems theory. Essentially he argued that it is a set of components necessary for a creative and productive organisation.

On the other hand, Senge's ideas are more than a mere checklist. His genius was to bring together essentially a system of theories (the 5 disciplines) which are themselves at least partly systemic. In the words of my colleague, Ken Wilson from WEB Research[#], he brought social science to cybernetics. To that I'd add psychology, cognitive, group dynamics and organisational change theories. So it's in.

What is "5th Discipline" ?

Essentially, Senge says that if you are to promote creative and humane institutions that are truly systemic in their approach, then you need to use 5 critical ways of understanding (or promoting) what is going on.

Systems Thinking
Essentially this is a repackaged form of system dynamics. Indeed some would argue that Senge rescued system dynamics from oblivion.

Mental models
"Mental models are conceptual structures in the mind that drive cognitive processes of understanding. They influence people's actions because they mould people's appreciation of what they see. People therefore observe selectively. Mental models most often invisibly define our relationship with each other and with the world in which we find ourselves …. Consequently mental models can undermine "systems thinking" by limiting the vision of what can be seen and done."[14] In other words, we see what we believe, not believe what we see. This "mental model" perspective draws heavily from the discipline of Action Science developed by two other Boston based academics Chris Argyris and Don Schön.

Shared vision
Shared visions create the energy and focus for successful endeavour. To use Ackoff's phrase it makes the system *purposeful* rather than merely *purposive*. On the other hand, Senge insists that it is not primarily what visions *are* that matter but what they *do* to people. By saying this, he essentially locates visioning in the cognitive psychology discipline rather than the "strategy" discipline.

Personal Mastery
Like many of Senge's components, this is more subtle than it appears to be. At first glance this emphasis on skill development sounds like a plea on behalf of the training and education industry. On closer inspection this is far from the case. Indeed he implies that these industries pose as much a threat to personal mastery as an opportunity. For Senge, personal mastery is a state of mind those who wish to gain full knowledge of themselves and the world around them; constantly deepening

[#] http://www.webresearch.co.nz

their own sense of personal vision. As Bob Flood says it is "a calling of intrinsic desires not a purpose to pursue"[15]

Team learning
In Senge's eyes, the purpose of a team is to create something greater than the sum of its individual members. This is achieved by creating a collective alignment between personal mastery and shared visions. The critical mechanisms Senge identifies for doing this are dialogue and discussion. According to Senge, the primary purpose of *discussion* is to win, whereas the primary purpose of *dialogue* is to understand; especially to understand the mental models of your colleagues. "In dialogue, a group explores complex, difficult issues from many points of view. Individuals suspend their assumptions but they communicate their assumptions freely…. In dialogue people become observers of their own thinking"[16]. Senge, here draws largely from the discipline of Bohmian dialogue, and some communication theory.

Strengths and Weaknesses of "5th Discipline"

Strengths
- It's popular and has entered beyond the "systems" world
- It links people with the more mechanical aspects of systems inquiry
- It stresses reflective processes, dialogue and experiential learning

Weaknesses
- It is powerfully driven on notions of consensus, which some feel reduces the potential for critical reflection
- In any formal sense, it does not address power, knowledge distribution and ethical issues
- Of itself it is not really a systems approach, it refers back to system dynamics for that. This can be constraining.
- Its whole is not particularly rigorous, even if the parts are.
- There is no formal means of boundary setting

Implications for evaluation of "5th Discipline"

With the possible exception of system dynamics, the core elements of the "5th Discipline approach are familiar to many evaluators and evaluations. It is in the detail where the differences lie. Evaluation is strongly focussed on "mental models" though evaluators may not consciously use the same theoretical base that Senge does. Many evaluators observe the impact of teams, but not distinguish between the relative contributions of dialogue and discussion. Training and skill development are at the core of many evaluations, but the more personal commitment to mastery isn't always the focus. Visions are frequently explored but the extent to which they are sharing by all stakeholders often isn't.

So the building blocks are familiar, but I suspect few evaluations are carried out with the particulars that Senge has in mind. I'd claim these omissions are critical.

The way to understand how critical, is to reframe Senge. Consider it not as a systems approach, but as a program theory. It is a systemic theory of how to create effective

and humane organisations. In which case, the sub-theories that make it up (eg system dynamics, action science, Bohmian dialogue) have to be part of the package. It's in these terms perhaps, that Senge is most useful to evaluators – as a model to compare reality with, than a methodological approach to inquiry.

CRITICAL SYSTEMIC THINKING

If Bertalanffy was the grandparent of current systems approaches, then Ackoff, and Churchman are the parents. They brought "thinking" into systems. However, Churchman brought something else, which is missing in many other approaches to systems thinking. He brought a strong sense of ethical and moral alertness, which dramatically alters the possibilities of systems thinking.

What is Critical Systemic Thinking ?

I can't improve much on Bob Flood's description, although I've done a bit of editing.[17]

"Churchman's systemic thinking is captured in four principles and seven central concepts. The principles are :-

- The systems approach begins when first you see the world through the eyes of another.
- The systems approach goes on to discover that every world-view is terribly restricted.
- There are no experts in the systems approach.
- The systems approach is not a bad idea.

The four principles are embodied in the following seven central concepts of Churchman's systemic thinking.

1. Purpose
Churchman set out nine conditions that must be fulfilled for 'a system' to demonstrate purposefulness. These conditions are as follows.

- A system is teleological.
- A system has a measure of performance.
- There is a client whose interests are served.
- A system has teleological components (ie the parts have purposes too).
- A system has an environment.[#]
- There is a decision-maker who can produce changes in the measure of performance of a system.
- There is a designer whose design of the system influences the decision maker, leading to changes in the measure of performance of the system.
- The designer aims to maximise the system's value to the client.
- There is a built-in guarantee that the purpose of the system defined by the measure of performance of the system can be achieved and secured.

2.Sweep-in
For Churchman, systems were not real entities existing 'out there' waiting to be identified. Rather, systems are whole system judgements, that is, judgements made in the knowledge of the totality of relevant conditions. This suggests, "sweeping-in" ever more features of the problem context. Sweep-in helps participants to become more aware and increasingly able to appreciate contrasting systems of meaning. It is

---

[#] Which places Churchman's ideas into the realm of so-called "open" systems, in contrast to "closed" systems which have no environment (eg some approaches to system dynamics)

an attempt to raise understanding, rather than to realise absolute knowledge. Sweep-in is a process of critical reflection that helps people think or debate.

3.Unfolding
Unfolding draws upon the nine conditions of system purposefulness. The conditions help people to add structure and meaning to their experiences. People may employ them to surface plausible interpretations of events. This aids them to identify possible clients, designers, decision makers and other affected people. It helps participants to consider measures of performance and who might gain or suffer from a particular design. Unfolding is therefore the critical counterpart to sweep-in. Sweep-in and unfolding culminate for Churchman in boundary setting.

4.Boundary setting
Defining an action area from the problem context through sweep-in and unfolding, centres on drawing boundaries around possible clients, and consequently surfacing issues and dilemmas relating to those clients for discussion. Boundary setting is an issue of great importance to critical systemic thinking. Put succinctly, the questions are, 'Who is embraced by the action area and thus benefits? Who is out and does not benefit? What are the possible consequences of this? And, how might we feel about that?' Boundary setting thus raises questions of ethics, efficiency and effectiveness, in a search for improvement and shows them to be inextricably linked. Boundaries are always open to further debate.[#]

5.Securing
According to Churchman, to secure improvement means that over time improvement persists, i.e. improvement must be sustainable. Systems of measurement are central to knowing that improvement has been secured. This means considering choices for improvement by critically taking into account and tracking long term environmental and developmental implications. The aim is to make choices that deliver ecologically viable in addition to socio-economically and socio-culturally desirable future improvements. So, the sweep-in process, for example, must embrace our children's future. Our children, some yet to be born, must be recognised as possible clients of today's decision processes.

6. Wisdom
Wisdom is thought combined with a concern for ethics.

7. Hope
Hope is the spiritual belief in an ethical future.

Strengths and weaknesses of Critical Systemic Thinking

Strengths

- It adds human principles and ethics to the mix
- It emphasises the essentially moral decisions and dilemmas associated with decisions made during an investigation and operational design

[#] Flood's own systems approach (see next section) goes even further. He states that every endeavour has first to decide a boundary. That decision is essentially political and ethical, not technical. He implies that evaluation, for instance, is inherently a political and ethical act not just because of the "value" content, but also because of the "boundary" decision.

- It stresses the importance of boundaries, and especially their ethical and political nature.

Weaknesses

- Being deeply political, it explores areas of systems that many just don't want go.
- The concepts and practice can be considered a bit esoteric
- The personal demands on the inquirers are very high

Implications for evaluation of Critical Systemic Thinking

At this stage there is some danger of repeating points already made. So what distinctive implications are there for evaluation of Churchman's ideas ? Being ethically alert and aware of the moral implications of what we are doing and creating are probably the strongest. Perhaps one of the things that distinguishes evaluators from other "inquirers" is that sense of moral responsibility. We are often walking around and occasionally treading on people's dreams. We judge or often promote judgement of those dreams. The implications of that are profound, and what Churchman's approach does, is give us some structure by which these moral dilemmas can be framed and worked through.

SYSTEMIC THINKING

The English, although somewhat peripatetic, organisational development thinker, Bob Flood recently published a major review of systems approaches that reassesses many of the ideas that underpin them.[18] Out of that flowed what he called the main components of "systemic thinking". I've included almost his entire final chapter, since I think it is not only an interesting systems approach in itself, but neatly sums up the complex area that systems based inquiry has become.

What is Systemic Thinking ?

The theory

Flood states that we work within three basic paradoxes :-

- We will not struggle to manage over things - we will manage within the unmanageable.
- We will not battle to organise the totality - we will organise within the unorganisable.
- We will not simply know things - but we will know of the unknowable.

He believes (probably correctly) that these three paradoxes of systemic thinking are mightily thought provoking. If they were embraced by human kind, Flood argues we would witness profound changes in the way we conceive ourselves as a species on planet Earth and the way we handle ourselves in everyday life.

How does he reach this conclusion ? This flows from his particular concept of systems, and how to manage them, explore them and ultimately evaluate them :-

- *Systemic awareness begins with a spiritual appreciation of wholeness*
- *Wholeness may be appreciated in terms of:*
  interrelatedness of events, and
  spontaneous self-organisation leading to emergence and new order.
- *Interrelatedness of events and spontaneous self-organisation in the natural sciences may be characterised in terms of:*
  deterministic and probabilistic feedback,
  operating according to laws that so far have withstood refutation,
  leading to emergence and new order.
- *Interrelatedness of events and spontaneous self-organisation in the social sciences may be characterised in terms of:*
  adaptive feedback,
  operating according to social rules and practices that people either wittingly or unwittingly agree upon,
  that people might desire to change by forming coalitions around an issue in spontaneous self-organisation,
  which may or may not lead to new order.
- *Human existence is intrinsically unknowable to the human mind.*
- *We know of some things, but only those which are local to us in space and time.*
  In space - things that we are immediately involved in (not simply in a geographical sense).
  In time- not very far into the future, or indeed the past.
- *Beyond what is local to us is unknowable.*
- *What we do know is a matter of interpretation and is mysterious.*

- *We live all aspects of our lives between mystery and mastery*
  There is mystery in the interpretive nature of what we know,
  and in the unknowable that lies beyond our interpretation, yet
  there may be a certain 'mastery' of the moment, and
  living between mystery and mastery means learning one's way into the future.
- *Mastery of the moment means that:*
  it is not possible to plan over a wide spread of interrelationships,
  through recurring emergence, far into the future; yet
  it is possible to learn about what is local to us, that is,
  it is possible to learn within the unknowable.
- *Learning a way into the future in 'management and organisation' involves:*
  explicit awareness of the bounded nature of our designs and decisions,
  explicit awareness of the ethical nature of our designs and decisions,
  appreciation of optional designs and decisions in the local context of the action area, and
  deepening systemic appreciation of issues and dilemmas that characterise designs and decisions.
- *Systemic thinking challenges the following concepts of traditional 'management and organisation':*
  problem,
  solution,
  normal organisational life,
  consensus,
  medium and long term plans, and
  prioritisation of ends over means.
- *If systemic thinking has its way, then:*
  we will not struggle to manage over things, we will manage within the unmanageable;
  we will not battle to organise the totality, we will organise within the unorganisable; and
  we will not simply know things, but we will know of the unknowable.
- *Systemic thinking in 'organisation and management' in essence is about being:*
  ethically alert,
  critically reflective,
  appreciating issues and dilemmas that we face,
  exploring possible choices for action, and hence
  may be known as critical systemic thinking.
- *We don't know very much about anything and actually never will.*

The practice

In method terms, Flood puts his money where his mouth is. He proposes a way that these principles and properties can be put into action. The data for this approach is based on four "windows" on the system :-

- Systems of processes
- Systems of structures
- Systems of meaning
- Systems of knowledge-power

This data is then incorporated through dialogue into a highly sophisticated and rigorous form of scenario planning and what he calls "systemic evaluation". Yes evaluation. He is one of the few systems thinkers to stress the importance of evaluation rather than (essentially) performance measurement.

Strengths and weaknesses of systemic thinking

Strengths

- It genuinely seeks to integrate some of the key systems approaches, without becoming a mess of compromises
- Flood has developed a method to operationalise his approach

Weaknesses

- Baldy stated, the principles and components do look like a shopping list of wishful thinking. There's quite a "bloody hell" factor.
- It's not at all obvious how to the ideas into practice.
- As yet there is no substantial community of practice for the approach

Implications for evaluation of systemic thinking

I've yet to get my mind fully around the implications of Flood's principles. However in terms of his description of practice, systemic thinking is the only systems approach that explicitly places evaluation at the core of the method. The approach to evaluation expressed in his book is rather crude, but the opportunity he has opened up for evaluators is considerable. Unfortunately, he links this to a particular approach to scenario planning, that most evaluators and clients would have neither the skills, nor perhaps the inclination to follow.

Given the inherent conservatism of much evaluation, and how far from the "Boston School" systemic thinking lies, I suspect it will be some time before Flood's potential contribution to can be fully explored and realised.

BOB WILLIAMS
http://users.actrix.co.nz/bobwill

April 2002

[1] Ackoff, R.L. (1971) "Towards a system of systems concepts", Management Science 17.11.
[2] Beishon J & Peters G. Eds (1976) Systems Behaviour, Open University Press.
[3] Senge P. M. (1992) The Fifth Discipline. Sydney. Random House
[4] Flood R. L. (1999). Rethinking the 5th Discipline. London : Routledge
[5] This is an edited and slightly modified version of Bob Dick's module in Areol, an on-line action research course. See http://www.uq.net.au/action_research/areol/areol-session13.html
[6] Checkland P. Systems Thinking, Systems Practice. Wiley 1999
[7] "A model of a whole entity; when applied to human activity, the model is characterized fundamentally in terms of *hierarchical* structure, *emergent properties, communication* and *control"* Ibid
[8] Ibid
[9] Ibid
[10] Bob Dick Op. Cit
[11] Eoyang G Personal. communication.
[12] Drawn heavily from Eoyang, G. (1997) Coping With Chaos Lagumo Corp
[13] Eoyang G.H, Burkas T.H "Evaluation in a Complex Adaptive System" (publication details unknown)
[14] Quoted from Flood R. L Op. Cit.
[15] Ibid
[16] Senge P. M. (1992) The Fifth Discipline. Sydney. Random House
[17] Flood. Op. Cit.
[18] Flood. Op Cit