HUMAN SYSTEMS
DYNAMICS INSTITUTE

Evaluating Performance in a CAS

July 5, 1998

Eoyang, G. & Berkas, T. (1999) Evaluating Performance in a Complex Adaptive System. In "Managing Complexity in Organizations." Lissack, M. & Gunz, H. (eds.) Westport, Connecticut: Quorum Books, an imprint of Greenwood Publishing Group, Inc.

Glenda H. Eoyang
Chaos Limited
50 East Golden Lake Rd
Circle Pines, MN 55014

Thomas H. Berkas, Ph.D.
Search Institute
700 3rd Avenue South Suite 210
Minneapolis, MN 55415-1138

Abstract

Evaluation is a central issue in all organizations. Many standard evaluation tools, techniques and methods rely on basic assumptions about linear organizational dynamics (predictability, low dimensionality, system closure, stability and equilibration). Some of these assumptions are not valid when a system enters the regime of a complex adaptive system (CAS). New strategies are required to evaluate complex adaptive human systems. New tools, techniques and methods must integrate assumptions about the dynamical and complex nature of human systems. This chapter summarizes the characteristics of CASs from an organizational perspective. It identifies properties of an evaluation system that are consistent with the nature of a CAS and describes tools and techniques that promise more effective evaluation. Finally, it outlines the emergent role of the evaluator in a complex environment.

Introduction

Individuals, programs and teams at all levels of an organization are expected to assess and report on their performance. Groups choose to evaluate performance for a variety of reasons. Evaluation data establish a foundation for continuous improvement and build frameworks for fact-based decision making. Such data establish individual and group accountability and support the effective use of resources. Organizations in education, non-profit public service, government and business recognize the need for effective formative and summative evaluation. Funders, participants, elected leaders, stakeholders and other constituencies expect organizations to be able to evaluate performance.

Most evaluation processes are based on performance against predicted goals. Increasingly institutions that are not able to provide such basic evaluative information risk losing the support of their funders and other stakeholders. Historically, evaluation programs were developed to work in organizations that were assumed to be closed, stable and predictable. And in many situations, linear, low-dimension evaluation systems provided adequate data to represent organizational performance approximately. Such evaluation approaches were close enough to meet the needs of organizations and their supporters.

To be effective, however, an evaluation program must match the dynamics of the system to which it is applied. Recent research in organizational management, behavior and psychology indicate that human systems behave as complex adaptive systems. Organizational systems that were once stable are moving outside the range of linear, predictable behaviors and entering into the regime of chaotic or complex

adaptive behavior. Prediction and controlled performance toward a goal, like those assumed by traditional evaluation methods, cannot be expected from a complex adaptive system (CAS). For this reason, new tools, techniques and methods must be designed to meet the needs of constituencies that request evaluation of organizations while they are in the more dynamic phases of complex adaptation.

Recent research reflects two approaches to applications of complexity in the evaluation of human systems. The first focuses on new descriptions of the evaluation process. It pulls metaphors from various CAS applications to explain the evaluation process as emergent and complex. As heuristic tools, these studies may encourage constructive conversations about the roles and strategies of program evaluation. They do not, however, offer integrated theory or pragmatic suggestions to improve the practice of evaluation. The second approach provides new measurement techniques to support data collection and analysis in complex systems. The most promising of these techniques is nonlinear time series analysis, which allows the evaluator to investigate patterns of behavior that emerge in a complex system over time.

This chapter seeks to build a third type of bridge between complexity science and evaluation of performance of human systems. By bringing together the complex and adaptive characteristics of human systems and the aims of evaluation, the chapter seeks to identify new theoretical grounding and new practical approaches that will improve the effectiveness of evaluation of human systems. It describes the characteristics of CASs that relate to evaluation design and implementation. It defines some principles that should guide evaluation in a CAS and introduces five tools and techniques to support evaluation in a CAS. Finally, it outlines the role of the evaluator in a complex environment. The purpose is to use complexity theory to bridge the gap between the assumptions of traditional, linear evaluation systems and the dynamical behavior of the human systems they are designed to assess.

Characteristic Behaviors of CAS

A CAS is defined in terms of its parts, the behavior of those parts and the emergent behavior of the whole. A complex adaptive system (CAS) consists of interdependent agents. The behavior of each agent conforms to a short list of simple rules, and the group of agents exhibits emergent, system-wide patterns of behavior.

All CASs share some features and behaviors in common. A subset of those characteristics, specifically ones that are relevant to evaluation, are addressed below. Each will be defined and its effect on evaluation systems will be described. These characteristics include:

- Dynamic
- Massively entangled
- Scale independent
- Transformative
- Emergent

Dynamic

A CAS exists in a state of dynamic flux. Because of the number of agents, their interdependence, and their openness to external influences, a CAS changes constantly and discontinuously.

Constant change in a CAS is driven by the number of agents, their association with their own rules of behavior and the interdependence between the agents and their environments. These complex interactions generate a system that is roiling with change. At no point does the system come to a natural equilibrium or stopping point. Many different metaphors have been used to describe this dynamic phenomenon in human systems. You can imagine such action to be permanent whitewater, a sand pile, shifting sands, unshackled action, coupled fitness landscapes, or any number of computer simulation models. All of these images connote the ever-changing nature of a CAS.

This change does not always follow a smooth, predictable pattern. Change happens at every point in time, but it may bring surprising outcomes. From a traditional point of view, any continuous change implies a smooth curve of effects over any given interval. This constraint does not affect the behavior of the CAS. While change in a CAS is continual (the system is always in motion), the change may not be continuous because it may not follow a smooth, predictable curve. It may come in bursts that are apparently random. Random jumps and discontinuities (bifurcations or punctuated equilibria) shape the emergent dynamics of the CAS.

For this reason, the evaluator cannot expect a smooth, linear path between project start and project end. System performance does not improve along a straight line or even a smooth curve. Bursts of activity may be preceded by long periods of apparent stasis. There may be no correlation between the percentage of time or resources consumed in a project and the percentage of distance moved toward a goal. Such unpredictable patterns cannot be assessed by means of periodic sampling or end-point evaluation only.

These temporal and dynamical characteristics challenge many of the assumptions of the traditional evaluator. The evaluator cannot realistically consider an organization or a program to be moving in a predictable way toward a pre-determined end point. This means that social systems do not move inexorably toward a project's end point. They may not come to rest even when the end of a project is reached. An evaluator may be able to assign an arbitrary beginning and end date of an intervention, but the system itself recognizes no such boundaries in time. For this reason, the whole concept of projected and predictable outcomes is an artificial construct when evaluating performance in a CAS. An evaluator may be able to frame expectations, but the self-organizing nature of the system may result in completely different outcomes than those expected.

Evaluators and evaluation plans must adjust to the perpetual but unpredictable dynamic behavior of a CAS. The changing patterns within the system must be captured and described, without depending on natural end points of behavior or extrapolation or interpolation from timed samples.

Dynamic Evaluation Principles. Because a CAS is dynamic, evaluation systems should incorporate flexible and dynamic features. Specifically, they should:

- Capture an emerging model of causal relationships. Assessment captures and describes change in a system based on current understanding of the context and causes for that change. Change can only be acknowledged when compared to some baseline or starting point. Because the patterns of causation are one of the things that change in a CAS, it is critical to capture baseline representation of those causal relationships, but it is necessary to revise that image frequently. Evolution of the causal model over time can provide a powerful and simple description of the systemic aspects of change.
- Evaluate and revise the evaluation design often. Because the CAS baseline is constantly shifting, the evaluation plan should include options for frequent and iterative reconsideration and redesign. Data about the redesign of the evaluation program can also become a rich source of information about the developing patterns of the system.

- Capture, preserve and learn from the "noise" in the system. Most evaluation programs focus on the relatively narrow range of expected behaviors. In a CAS, much of the meaningful information about system future, patterns and dynamics come instead from the unexpected system behaviors. For this reason, evaluation should capture the unexpected as well as the expected, the long- and short-term outcomes and the close and distant points of view. Only from this diverse data can an evaluation emerge that is sophisticated enough to reflect the complexity of the system being evaluated.

Dynamic Evaluation Tool. In the dynamics of a CAS, even the relationships of causes and effects change over time. As described above, for maximum utility, an evaluation approach needs to capture the changes in causation during the course of an intervention. Figure 1 gives one example of such a causal diagram, which is one option for representing the causal logic of a system. A variety of other symbol sets would work as well. Other potentially useful methods include stocks and flows modeling from system dynamics, process modeling from information systems design techniques or mind mapping from creativity methodologies.

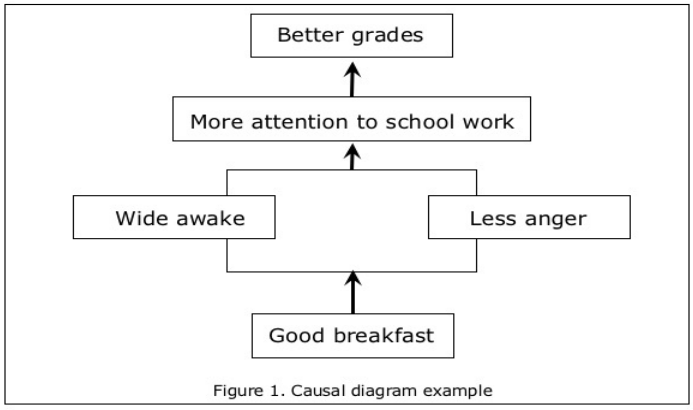

Figure 1. Causal diagram example

Regardless of the systems used, to be most effective the representation should be:

- Graphic and well labeled
- As simple as possible
- Generated by a group of concerned and involved stakeholders
- Widely distributed for use throughout the system
- Reviewed and revised frequently

A causal diagram provides many different benefits to an evaluation program. Conversation about basic cause and effect reasoning in the system will help surface and reconcile divergent mental models. The causal diagram can establish the foundation for other aspects of the evaluation system by defining major variables, indicators and

hoped-for outcomes. Periodic review and revision provide natural points of evaluation and assessment. Major shifts in the causal diagram signal discontinuities in the dynamically changing CAS environment.

Massively Entangled

Relationships in CASs are complicated and enmeshed. Kontopolous (1993) describes CASs as massively entangled because the component parts of the systems and the variables describing those parts are large in number and interrelated in complicated ways. Two kinds of entanglement relate directly to evaluation: among the large number of variables that determine system behavior and among system participants.

Many CASs are driven by a large number of interdependent variables. The behavior of most CASs is influenced by a wide variety of factors. (The exception, of course, is deterministic chaos, in which system behavior emerges from a small number of nonlinear relationships.) In addition to being numerous, variables can be nonlinear and discontinuous. Some dimensions vary in their influence over time. They may lie dormant for long periods until some control parameter reaches a critical value and sparks them into action.

One way to respond to these multiple and unruly variables is to look for those "differences that make a difference" in the system. The number of potential differences is quite large and virtually unknowable in social systems. For example, adolescents' behavior may depend on age, gender, physical type, family stability, relationships with peers and so on. In addition, over time, the relevance, power or interrelationships among differences may change. Such a complex interaction of variables makes it unrealistic to expect to represent the system with a finite number of independent and dependent variables. Though some current trends in psychological assessment respond to multidimensionality of system response, most traditional evaluation systems seek to identify a small number of key variables that affect change and to establish the relationships among those variables. Unfortunately, these strategies are insufficient to represent the complicated interdependencies in a CAS.

In addition to being intractable (if not infinite) in number, the dimensions that drive behavior of a CAS have nonlinear relationships with each other. A small change in one variable may generate exponential change in another. This pattern is exacerbated because CASs depend on iterative processes. They repeat the same processes over and over again. The output of a previous process becomes the input for the next one. Iteration magnifies the effects of the nonlinearity, so that simple causal relationships are virtually impossible to detect, to measure, control or evaluate.

These complex interrelationships demonstrate that the entanglement of variables in a CAS involves both their number and nonlinearity. These complex relationships limit the value of traditional causal reasoning, making it difficult to do at best; counter-productive at worst.

Another level of entanglement, however, introduces a different set of complexities to assessing behavior in a CAS. The parts of the system (individuals, groups, institutions) are related to each other in complicated and unpredictable ways. These systems demonstrate cross-current causality, which complicates analysis. Kontopoulos describes this characteristic of CASs as heterarchy, saying that their structures lie somewhere between the extremes of anarchy and hierarchy.

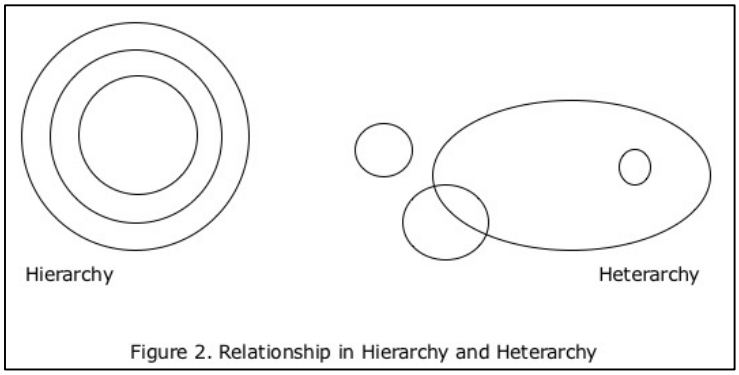

Figure 2. Relationship in Hierarchy and Heterarchy

In Kontopoulos's model, anarchy allows only for causality from the part to the whole. Individuals are free to take action, and their actions determine the behavior of the whole group. Hierarchy, on the other hand, assumes only top-down causality. (Whatever is determined at the top of the organization will be carried out in all its parts.) Neither of these alternatives is rich enough to describe the range of behaviors of the CAS, whose interactions are driven by multi-directional causality. Sometimes the part determines the whole.

Sometimes the whole determines the part. Sometimes, parts determine each other. In a social system, for example, an unruly individual may necessitate a change in the rules, or pressure from peers may be sufficient to constrain the behavior of this individual. Each level of organization has a definite, though incomplete, causal relationship with all others.

In addition, the parts of the system are not related in simple, logical patterns. The structure of the system is not a simple arrangement of circles in a concentric set, with each higher level encompassing the lower ones completely. Rather, the system performs as a set of non-concentric, interlocking spheres of influence. For example, in a simple pattern a child might be described as a member of a class within a grade level in a school in a district in a geographical area. In this logical, concentric model, each of these sets is totally encompassed in the next larger one. Though such a description is possible, it overlooks the rich interdependencies of other heterarchical relationships of family, friendships, religious associations, neighborhoods, reading groups, scout troops and so on. Such messy and multi-level causal relationships determine the behavior in a CAS and make it difficult for the evaluator to establish clear units of analysis or lines of causality.

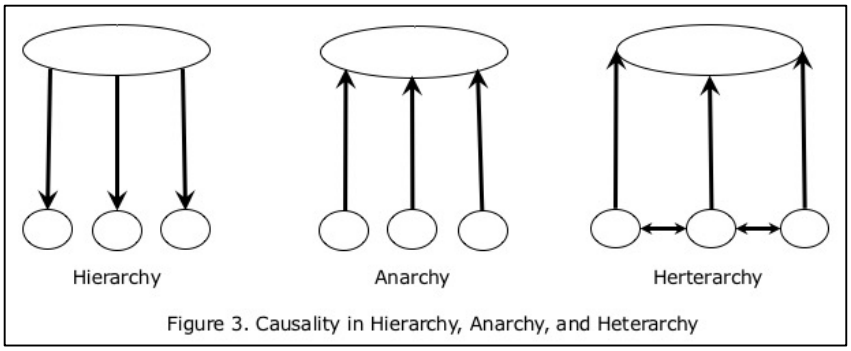

Figure 3. Causality in Hierarchy, Anarchy, and Heterarchy

Most traditional evaluation methods, both quantitative and qualitative, make assumptions to constrain the system and build a simple model of cross-scale relationship. Some assume that the system is based on concentric levels of organization to avoid the issue of complex crossing of boundaries between organizational levels. The evaluator may define the individual participant as part of an age cohort, which is part of a treatment group, which is part of a program and so on. Each part of the system has its mutually exclusive definition with regard to all the rest.

Other evaluation methods focus on a single organizational level to bypass the issues of heterarchy altogether. These methods choose either the individual or the group as the unit of analysis, and ignore the other levels to avoid cross-scale concerns and patterns.

Unidirectional causality is also a fundamental assumption for most evaluation programs. This approach identifies a simple, causal pattern as fundamental to its work. An evaluator will define a small set of causes and measure effects that are perceived to follow from those causes. Such a focus on pre-determined causality denies the essential multi-causal patterns in a heterarchical system. When an evaluator identifies variables and describes them as having dependent or independent relationships, the cross-causal nature of the system is ignored.

In a CAS, these simplifying assumptions are not valid. Because a CAS is heterarchical, linear methods of problem definition or methodological simplification are inadequate.

Entangled Evaluation Principles. CASs involve multiple "differences that make a difference" and complex interrelationships among system components. For this reason, evaluation systems should:

- Incorporate multiple strategies, cycle times, time horizons, dimensions and informants. Because a CAS has a structure that is nonlinear, open and high-dimensional, an evaluation design cannot pre-determine all factors that will be of interest. For this reason, it is critical that a variety of data be collected to reflect the variability of the system. The practice of triangulation, which is common in qualitative research methods, is an example of such a strategy. Triangulation of informants, strategies and

- Timeframes will help the evaluation program represent the complex dynamics of the system better. Such a diverse design will allow the evaluator to collect a wide range of information and to determine in hindsight what was most relevant. An effective CAS evaluation design will even incorporate linear evaluation strategies in short time frames and closed parts of the system where prediction appears to be possible.
- Be explicit about the language and meanings of evaluation findings. Interpretation in a CAS depends intimately on its heterarchical context. Evaluation will be distorted when the method is designed in one context, the data collected in another and findings reported in still another. The complex interdependencies and unique dynamics of each environment require that context-specific identifying information be associated with all data and analysis that are used for evaluation.

Entangled Evaluation Tool. Iterative redesign generates an evaluation program that reflects the massive entanglements of the system. It also allows the evaluation design to co-evolve with the system under investigation. Iterative redesign establishes a plan for periodic review and revision of the evaluation design throughout the life of the project. The practical problem with iterative redesign is that it is time consuming and difficult to document and manage. This section suggests an approach that simplifies the process without losing its critical self-reflective advantages. It focuses on time horizons and collection/reporting scales of the system, which are the two main "differences that make a difference" in a dynamical system. It consists of a schedule of planned evaluations and a structured process to be repeated at regular intervals during the project.

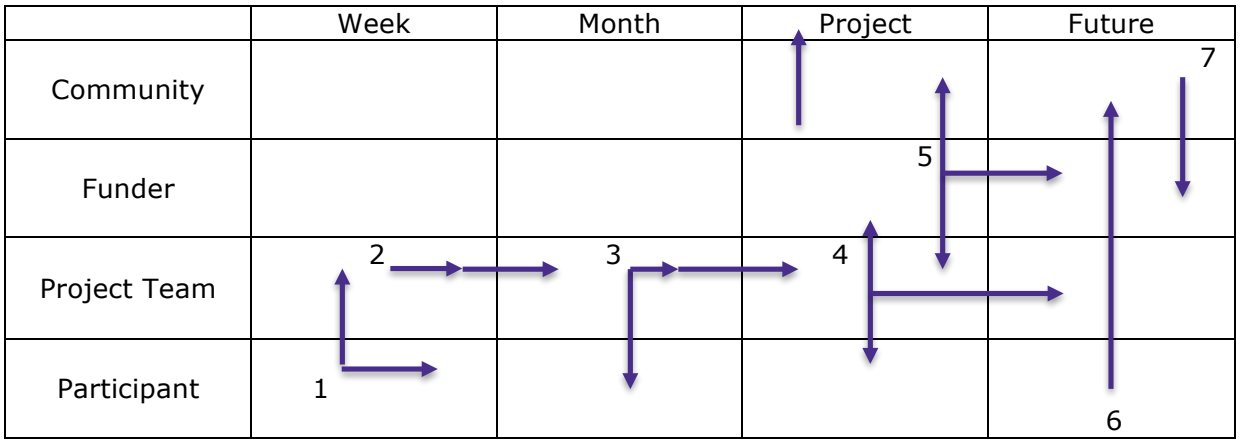

Figure 4. Iterative redesign example.

The structured process consists of a small number of questions to be addressed by each group at each numbered point in the schedule. At points designated on the plan, individuals or groups determine goals, evaluate performance against previously defined goals and review projections for outcomes for each future time period by responding to the following questions:

- How did we perform against the projected outcomes for this time frame?
- What outcomes do we project for the next iteration of each time frame (indicators, data collection and analysis methods)?
- What action should we take to move toward those outcomes?
- To whom and how should we communicate the results of this discussion?

A sample schedule of evaluation activities appears in Figure 4. Any particular project team will design the schedule to meet its needs. Because the points at which data are collected, analyzed and reported will vary from project to project, the arrangement presented here is illustrative only. It is based on the following assumptions:

- The near-term timeframes (week, month) can be varied to match the natural cycle time for the project. Some project outcome and long-term outcome frames should always be included, however. At each of these regular intervals, the evaluation design and performance against plan will be reviewed and revised. The shorter cycle times would represent more informal and the longer times more formal analysis of the plan. For example, the weekly redesign might consist of a five-minute discussion in a staff meeting, while the project review may involve a formal evaluation document.
- The arrows represent the collection and reporting patterns. Data is collected at the tail of the arrow, and findings are reported at the arrow's head. The thoroughness and formality of each of the data collection and distribution cycles would be determined by the team before the beginning of the intervention and modified as needed.
- The numbers indicate how many different evaluation contexts are included in the plan. This number will change with the needs of individual projects.
- The four levels of stakeholders (participant, project team, funder and community) represent the scaled nature of the CAS. The names and the number of the groups can be varied as needed for the specific evaluation context. These groups can even be intersecting and entangled if the complexity of the project demands it.

This process involves an iterative schedule of evaluation and redesign and a simple set of questions to be used in each iteration. Such a structure can ensure that intended outcomes will be defined for each time period and that performance will be measured against those outcomes. The approach, however, does not lock the system into an unrealistic expectation for prediction or control, which is unrealistic in a CAS.

Scale Independent

A CAS functions simultaneously at many different levels, or scales, of organization. Complex systems gain coherence across levels because the same patterns appear at various levels. For example, the angles between the veins in a leaf are the same as those between twigs, branches and roots. This self-similar structuring provides an integrity and comprehensibility to the system. The same scale independent patterning can be observed in human systems. Individual agents take relatively independent actions; various groupings of agents emerge in the dynamical course of events; and the whole system exhibits identifiable behaviors. Behaviors in each of these domains are both similar to and different from behaviors in the other domains. An individual child may reflect the tension felt in an entire school, groups or gangs may form in response to this tension, and rumors may move through the faculty in response to specific incidents. Each of these domains is intimately associated with the others and exhibits both similarities and differences from them.

The complex outcome behaviors of a CAS may be the result of the iterative application of a "short list of simple rules." A cellular automaton, which is a type of computer simulation model, demonstrates most clearly the effect of a "short list." In a cellular automaton, a collection of interdependent agents are generated and represented as patches of light or darkness on a computer monitor. The agents can change their state by turning their light on or off. Each agent watches its neighbors and follows a short list of simple rules to determine its own local behavior. Over time, collections of these interdependent agents mimic a variety of complex behavior patterns (including the flight of a flock of birds, the life cycle of a bacterial colony or the spread of infectious disease). The short list of simple rules is one mechanism that connects the parts to each other and to the whole and brings the coherence of scaling to the otherwise apparently order-less behavior of a CAS.

Traditional evaluation methods make basic assumptions about the relationships between the part and the whole in a human system. They do not depend on patterns that appear at multiple scales nor on the possibility that complex patterns of behavior emerge from short lists of simple rules. Quantitative approaches to evaluation assume that uncontrolled interdependence among participants is minimal. The behavior of the whole group is seen as the sum of the behaviors of its parts. The co-evolving, scale independent nature of the CAS makes such an assumption unrealistic. In some CAS situations, the individual can act as proxy for the whole because certain patterns are constant across scales of the system. When behavior is driven by a short list of simple

rules, observations of an individual provide clues to the behavior of the whole. In other situations, the emergent patterns are different in kind from the summation of the parts. New and unexpected system-wide behaviors emerge from the complex interactions of the agents. Traditional evaluation systems are not designed to deal with self-similarity or the radical emergence that are evident in scaling phenomena of CASs.

Scaling Evaluation Principles. Because a CAS incorporates many self-similar levels of organization, an evaluation program must incorporate both micro- and macro-patterns and structures. A specific evaluation program should:

- Make information about the evaluation process open and accessible to all stakeholders. When information moves smoothly between and among organizational levels, self-similarity and its resultant coherence are encouraged. By being explicit about decisions and processes, evaluation becomes an effective transforming feedback loop. As a process with multiple stakeholders, evaluation can support change and transformation efforts by participants in all parts of the system. Used in this way, evaluation becomes a part of the intervention, rather than some irrelevant activity.
- Be sensitive to both the similarities and the differences between contexts within the same system. Design and develop evaluation processes at the level where they will be used. Because certain system patterns are independent of scale, system-wide evaluation will uncover system-wide patterns. General rules and short lists of simple rules can be developed to apply to the entire system. On the other hand, each context has its own unique context, so specific evaluation plans must be designed to meet the unique needs of the local context. Not only will the measures be more relevant and meaningful, but also the process of defining the evaluation plan will contribute to the system transformation process.

Scaling Evaluation Tool. A short list of simple rules gives coherence across scales of a complex system. Given this dynamic of a CAS, it is possible to develop a short list of simple rules that could generate a complex and effective evaluation program across many different parts of a complex human system. Such a strategy would use the inherent dynamics of the CAS to establish an evaluation program. This approach would involve developing a short list of simple rules that will help each individual and group in the system design and implement their own evaluation plans.

The system-wide evaluation plan that would emerge from this process would not be predictable because it would evolve as the system evolved. This approach provides a practical framework for the constructs of micro-design and micro-evaluation. Micro-design and evaluation acknowledge that the system is changing too quickly to support large-scale planning or assessment. As an alternative, designers and evaluators focus on the smallest stable element (time, organizational unit, or functional component). An intervention or assessment is designed for each small unit. A consistent short list of simple rules would provide some coherence at the macro-level as a foundation for mico-

design. Such an evaluation system would emerge and integrate with the on-going evolution of the system under observation.

Every part of the system would be expected to follow the same short list of simple rules for evaluation. The following rules might be sufficient to establish such a reflective evaluation process:

- Evaluate to inform action.
- Communicate findings to others in terms they care about and understand.
- Focus on "differences that make a difference."

If all stakeholders of a program followed these three rules, they would generate a cluster of evaluation activities that would look different than many traditional evaluation plans. One certainly could not predict the activities or the findings of the evaluation process, but all participants in the CAS would be informed productively about their performance and the performance of others in the system. The role of the professional or external evaluator in such a situation would be to help everyone involved watch the evaluation program emerge and to follow the same short list of simple rules to assess and improve the emerging evaluation process as a whole.

Transformative

The interdependent agents of a CAS are transformed and transforming in their interactions. Because a CAS and its agents are open systems, transformation occurs across the system's external boundaries. Feedback loops generate both change and stability in the system. Because each CAS is unique its behavior is context dependent. All of these transformative behaviors have major implications for the design and implementation of effective evaluation systems.

The agents that constitute a CAS are indefinite in number and are acted upon by forces external to the system. Bertalanffy, in his seminal work ***General System Theory: Foundations, Development and Application***, acknowledged that open systems were different in kind from the closed, well behaved systems that were his focus. Weick applied the ideas of open systems and unplanned interdependencies to articulate a model of organizational behavior. Open systems are unpredictable, and their behaviors are dependent on context. System boundaries must be defined arbitrarily, and factors outside those boundaries may have as much influence on system behaviors as the dependent and independent variables defined within it. Agents and causal factors that lie outside the system at one time or place may be an integral part of the system at others. Because these systems are dependent on context and because each context is unique, CASs are themselves each unique. Two apparently similar systems may demonstrate profound differences over time. Even the same system, after the passage of time, may bear little resemblance to its previous configuration.

The transformative nature of a CAS influences evaluation in three ways. First, the evaluator cannot identify with any confidence which factors will influence outcomes. What appears to be relevant may become irrelevant, and the accidental may become causal in the course of an intervention. Second, participants in a system cannot be identified with any level of confidence. Sources of evaluation data may be inaccessible as they move in and out of the system of focus. While this feature complicates process evaluation, it makes longitudinal, individual study designs unfeasible. Finally, an evaluation system must be adaptable to the unique situation of each individual system. No matter how well a generic evaluation process is designed, it will be ineffective unless it is adapted to the unique situation of each local system under evaluation. A CAS cannot be evaluated in isolation from the environment in which it is embedded.

In the absence of a rigid external boundary, agents in a CAS are connected to each other by a complex network of transforming feedback loops. These loops carry resources (material, information and energy) from one agent to another. When an agent receives a resource, it adapts and sends out responding messages to other agents in the system. These transforming feedback loops serve to give both stability and changeability to the CAS. They fuel the interdependence of the system by keeping the parts synchronized. They support evolution of the system by providing impetus and resources for adaptation.

Feedback loops relate to evaluation in three distinct ways. First, evaluation is a powerful feedback loop. Designing an evaluation system and then collecting, analyzing and reporting findings generates a tremendous amount of potentially transforming information. This information may generate a variety of change, especially if it is accessible to individual agents, emerging groups and the system as a whole. The second way that feedback loops affect evaluation is as an object of evaluative focus. If feedback loops are the mechanism of transformation, it is logical to evaluate their effectiveness as a way to assess the performance of the organization as a whole. Finally, the evaluation process introduces a new set of feedback relationships in the system. By collecting and analyzing data, the evaluation process itself becomes a player in the complex emergence of behavior in a CAS. An evaluator must acknowledge this intimate interaction with the system.

Transformative Evaluation Principles. The CAS transforms and is transformed over time. Evolutionary change can be observed in individual and system-wide behavior over the course of an assessment period. Effective evaluation systems respond to this concern because they:

- Make evaluation a part of the intervention. As a transforming feedback loop, assessment activities should enrich and enhance the intervention activities. To support this goal, the evaluation design should be as simple and self-documenting as possible. It should include simple, iterative activities, and it should be totally understood by as many stakeholders as possible.

- Involve as many members of the system as possible in the design of the evaluation system. Because each CAS is unique, any effective evaluation system will need to be adapted *in situ*. The evaluator increases understanding of the system and its dynamics when changes in the evaluation program are made during the design stage and before implementation begins. Discussion of the design can be a time of tremendous learning and adaptation on the part of all stakeholders. With effective and continuing feedback, the whole system can co-evolve and adapt to the needs and aspirations of participants and the community.
- Use evaluation as a reinforcing, rather than damping, feedback mechanism. Especially early in a project, use evaluation procedures to find things to celebrate. Use the feedback to amplify the energy and commitment in the system.

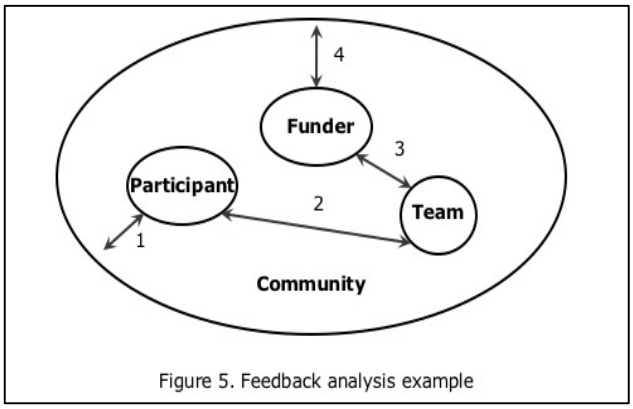

Figure 5. Feedback analysis example

Transformative Evaluation Tool. This approach represents an effort to identify and simplify the transforming feedback patterns for the evaluation process. It may also be used to identify the feedback loops that are intended to be part of the intervention. The entities in the system are represented in a Venn diagram. Lines between the components indicate the exchange of information, energy, and/or resources. Each arrow can be described in terms of:

- What information, energy or other resource flows each way along the arrow?
- What transformation can be observed as a result of the exchange?
- How might the feedback loop be adapted to be more transforming?
- Should new feedback loops be added to the system? If so, where and how?
- What should be included in reflective feedback loops (not shown in the diagram) within each of the component parts?

Each feedback loop, then, becomes a focal point for observation, measurement, assessment and intervention in the system.

Emergent

CASs exhibit emergent, or self-organizing, behavior. New patterns are generated by the interaction of the agents. New structures are established, and old ones disappear. These structural changes are not designed and imposed by some force outside of the system. They self-organize as the internal dynamics of the system play out over time. Two aspects of emergence in a CAS are of particular interest to evaluators: sensitive dependence on initial conditions and attractors regimes.

A CAS is sensitive to small changes in initial conditions. An apparently trivial difference in the beginning state of the system can result in enormously different outcomes. This phenomenon is sometimes called the "butterfly effect."

The implications of the butterfly effect for evaluation in social CASs are many. Traditional evaluation tools assume the effectiveness of an intervention to be a direct result of the intervention itself. In a CAS, however, the results of the same intervention may vary widely because of small variations in initial conditions.

One of the implications of sensitive dependence is a need for a shift in thinking about outcome evaluation. Traditionally, outcome evaluation depended on a prediction of outcome, behavior required to move toward the outcome and the measurement of progress to determine success or failure. Because of the butterfly effect, such a theory of "outcome evaluation" is not feasible in a CAS. Prediction is not possible and controlled performance toward a goal is unrealistic, so evaluation based on performance against such a goal is meaningless. It is feasible to define an outcome in a CAS when it is recognized as a possible scenario result rather than a predicted outcome. It is perfectly reasonable to have hopes for the future of a CAS, to take action in each subsequent moment in the context of that hope, then to evaluate whether the hope became a reality in a given amount of time. This CAS approach to outcome evaluation, however, does not imply the model of prediction and control assumed by many traditional outcomes evaluators.

In addition to concerns about outcome definition and evaluation, many common assessment techniques use reliability as a measure of quality. Reliability requires that the same evaluation process be used multiple times with the same results. Because a CAS is sensitive to initial conditions, however, it is impossible to reproduce the same evaluation environment twice. For this reason, reliability is not a logical possibility in a CAS, so some other criterion is required to assess the effectiveness of evaluation programs.

Because of the butterfly effect, the future of a CAS is unpredictable, but not all of the future is equally uncertain. The near-term future can be relatively knowable; mid-term will be less predictable; and far-term future will be uncertain at best. As the time horizon moves out, uncertainty increases because there will be more opportunities for various conditions and more time for those variations to magnify uncertainty.

Because of this dynamic, an evaluation program must have distinct strategies for assessing a range of possible short-, medium- and long-term goals.

Sensitive dependence is one aspect of emergence in CASs, and it drives concerns about outcomes, reliability and time horizons for evaluation systems. A second aspect of emergence deals with patterns that appear over time in the behavior of CASs. These patterns, called "attractors," provide some insight into the emerging relationships in a CAS.

Systems exhibit certain classical patterns of behavior over time. Scientists describe these patterns as attractors. An attractor is not like a magnet--it does not literally constrain the behavior of individuals in the system. Instead, an attractor is the pattern that forms as the individuals in the system interact. The individual behaviors form the pattern, and then other individuals are constrained to perform within the pattern. In this way, an attractor is emergent and self-reinforcing. Wherever an agent enters the system, it will move toward the established pattern of behavior, which is the predominant attractor regime. In this way, the activities of the agents in a complex system are patterned, though they are not predictable for any specific individual at any particular time.

Systems in motion generate one of four distinct patterns. The four categories of attractor are generally recognized to be point, periodic, strange and random. A point attractor emerges when all parts of the system tend to converge to a single point. The common example of a point attractor is a marble rolling around in a round-bottomed bowl. Wherever the marble starts out, it will eventually reach the lowest point in the bowl. A periodic attractor emerges when the system oscillates from one value or position to another. Examples include a driven pendulum, the swimming suit industry, or the monthly financial reporting cycle. In all of these cases, the cycle is defined and the system moves from one orderly position to the next. A strange attractor, which is characteristic of the behavior of a deterministic chaotic system, represents system behavior that stays within a bounded region without ever repeating the same sequence twice. A random attractor shows no discernible pattern at all. Each one of the attractor regimes describes an emergent pattern of behavior that is exhibited by a system, even though the behavior of the individual agents is unpredictable from one moment to another. By collecting time series data, the evaluator can reconstruct a portrait of the attractor regime for a particular system. The attractor is the primary method of "seeing" system-wide changes in behavior over time.

Studies of system maturity and creativity indicate that as systems mature, they tend to move from one attractor regime to another, beginning with random and moving through periodic, strange and point attractor regimes in sequence. Discerning systemic patterns of behavior through attractors should be a standard technique for evaluation. The problem is that traditional techniques of evaluation do not provide the kind of data required for reconstruction of the system attractor. Attractor patterns can be discerned

only from carefully designed time series analysis. The series must be of sufficient length. The sampling interval must be constant and scaled to the dynamical behavior of the system. Traditionally, many evaluation efforts only collect data at long and discrete intervals, such as the beginning and ending points of an intervention. These techniques do not capture the emerging patterns over time, so they are unable to reconstruct the attractor portrait for the system. Because the emerging attractor is the most trustworthy picture of system- wide behavior, evaluation methods should be designed to ensure that they capture and analyze the data that would reveal such patterns.

Emergent Evaluation Principles. The system-wide behaviors of a CAS emerge over time. For this reason, the evaluation system should focus on developmental and emergent patterns of behavior that:

- Match the developmental stage of the system. Some evaluators refer to this as readiness assessment or evaluability assessment. They ask the question, "Is this system mature enough to provide useful data during an evaluation?" A CAS working in the regime of a random attractor may not provide information about system-wide progress toward outcomes, even though individual agents are progressing well. Such a system will require different evaluation techniques than a system that has matured to a point attractor regime. Consider the dynamical pattern that the system exhibits over time to design an evaluation program to capture the "differences that make a difference."
- Track patterns and pattern changes over time, rather than focusing exclusively on behaviors of specific individuals or groups. While it may be unreasonable to expect a particular path of development or a pre-determined outcome from a CAS, emergent patterns of behavior can be expected outcomes. An effective evaluation system must be able to capture and report on these evolving patterns.

Emergent Evaluation Tool. The emergent nature of a CAS unfolds over time, so the only way to observe this emergence is through the use of time series analysis. In time series analysis, quantitative data is collected at regular intervals over a period of time. The sequence of numbers is analyzed to determine what patterns emerged in the data. A variety of analysis and modeling techniques can be used to discern patterns in time series data. Both linear and nonlinear techniques are available, and valuable information can be collected from both. Time series analysis has been used to assess various human systems behaviors in business, government and industry. This approach has also been used to evaluate psychological behavior in humans.

This approach has many benefits as a method to identify and describe the emerging patterns of behavior in a CAS, but it also has its drawbacks. Specifically, the approach:

- Requires a long time series for analysis. It is difficult to generate a time series of sufficient length from a human system.\ Is very sensitive to sampling rates. The sampling interval must be small enough to reveal the underlying pattern but not frequent enough to introduce irrelevant noise. Often, a sampling rate that is frequent enough for analysis makes unrealistic demands on the participants to collect and report data.
- Requires experienced analysts. Given the current technology, time series analysis is as much an art as a science. Analysts must have extensive experience with the tools and some familiarity with the context that generated the data before they can interpret the data realistically.
- Depends on mathematical algorithms that may not be valid or reliable. The tools for time series analysis are relatively new and are based on a variety of assumptions about the stability, stationarity and distribution of the system under study. They are not appropriate for all applications with human CASs.
- Works best on systems with few dimensions. Current tools are unable to distinguish between a very high-dimension attractor and randomness. Because most human CASs exist in high-dimension space, the tools may be ineffective.

In spite of its drawbacks, time series analysis and reconstruction of attractors promises to give real insight into the complex, entangled dynamics of evaluation of human systems. All of these tools and techniques (causal diagrams, iterative redesign, shorts and simples, feedback analysis and time series analysis) provide ways for the evaluator to capture and interpret information about the performance of a human CAS. When these approaches are used in conjunction with more traditional quantitative and qualitative evaluation methods, it will be possible to generate an assessment of a CAS that matches the variety and richness of the system itself. Such an assessment will yield information that can be used by all participants of the system to improve performance, even when prediction and control are not possible.

This section has presented a summary of CAS behaviors that are most closely related to evaluation. As CASs, human systems are dynamic, entangled, scale independent, transformative and emergent. These characteristics challenge the basic assumptions of traditional evaluation methods. They necessitate new evaluation approaches that are as rich and varied as the human systems they are designed to assess. The next section suggests a new role for the evaluator of such a dynamic system.

Role of the Evaluator

Complex adaptive dynamics do more than just require new tools and techniques for evaluation. They also transform the evaluator's role. Rather than being concerned with defining and measuring performance against specific outcomes, the evaluator takes on the task of designing and implementing transforming feedback loops across the entire system. This role of transforming agent falls into two primary categories: absorbing uncertainty and making learning the primary outcome.

Absorbing uncertainty. Change unfolds continually in a CAS. Individuals and their organizations express anxiety during times of change and uncertainty. Evaluators have an opportunity to mediate this anxiety in three ways. They can help the system understand and make sense of the CAS dynamics they observe. By explaining the basics of CAS, the evaluator can help the organization be reflective about their experiences and their fears. Second, evaluators can help articulate the CAS dynamics within a given, local context. By stating, and encouraging others to state, the dynamic patterns in the environment, the group can begin to build mechanisms to cope in the future. Finally, the evaluator can lower the cost of failure. By framing an evaluation method as experimentation and learning, the evaluator can encourage individuals and groups to value their mistakes and to learn from them. All of these intervention approaches can help the evaluator absorb the pain of uncertainty and lack of control that are hallmarks of the CAS.

Making learning the primary outcome. Effective adaptation is the best indicator of success in a CAS. Evaluators can use their experiences and expertise to focus on learning as an adaptive mechanism. To do this, the evaluator should:

- Emphasize the importance of variety in a system
- Distinguish between exploitative and exploratory learning and help groups find the appropriate uses of each
- Encourage the use of scenario planning and creative approaches to planning for the future.
- Encourage every individual in the organization to think about the multiple, heterarchical levels of organizational structure

From this perspective, the evaluator is less an instrument of assessment in the organization and more an instrument of transformative change. Evaluators can provide a valuable service to the CAS organization by designing, implementing and maintaining effective feedback loops between and among system components.

Conclusion

Characteristic Behaviors of CAS	Principles to Guide Evaluation in a CAS	Tools and Techniques for Evaluation of a CAS
Dynamic	▶ Capture an emerging model of causal relationships ▶ Evaluate and revise design often ▶ Capture, preserve and learn from noise in system	Causal diagrams
Massively entangled	▶ Use multiple strategies, cycle times, horizons, dimensions, informants ▶ Be explicit about findings and meanings of findings in various contexts	Iterative redesign
Scale independent	▶ Make information available to all stakeholders ▶ Design and develop the evaluation plan where it will be used	Shorts and simples
Transformative	▶ Make evaluation part of the intervention ▶ Involve participants in design ▶ Use evaluation to reinforce ▶ Co-evolve the evaluation plan with key stakeholders	Feedback analysis
Emergent	▶ Match developmental stage of the system ▶ Track patterns of change over time	Time series analysis

Figure 6. Summary of the behaviors, principles and tools and techniques for evaluation in CAS

As long as human systems behaved in linear and predictable patterns, traditional methods of evaluation were sufficient. As human systems move toward complex adaptive behavior, however, the assumptions that are the foundation for evaluation are no longer valid. In some circumstances, the traditional evaluation approaches are effective, and in others they are woefully lacking. A CAS perspective on evaluation opens the door to approaches that truly reflect the complexity and adaptation of the human systems they represent. Such an approach integrates divergent techniques of the past and presents new alternatives for the future.

As described in this chapter, a CAS approach to evaluation does not replace more traditional approaches. Rather, it provides a theoretical framework that incorporates many tools and techniques that were previously considered to be at odds with each other. Qualitative and quantitative approaches of all kinds can play significant roles in the evaluation of a complex adaptive system. A CAS model of evaluation is most useful when complexity renders other methods of evaluation ineffective; when evaluation will be used to challenge existing assumptions of linear causality; or when the interventions to be evaluated are designed to reflect the complex adaptive nature of the system.

This chapter has outlined the behaviors of CAS and related those behaviors to the issues of evaluation. It has also identified principles and tools for an evaluation program that would be effective in a CAS. Finally, it has outlined the role of the evaluator in a healthy CAS.

References

Bak, P. (1996). How Nature Works: The Science of Self-Organized Criticality. New York: Copernicus.

Bertalanffy, L. (1968). General System Theory: Foundations, Development, Applications. New York: George Braziller, Inc.

Briggs, J. & F. D. Peat (1989). Turbulent Mirror: An Illustrated Guide to Chaos Theory and the Science of Wholeness. New York: Harper Row.

Cooperrider, D. & S. Srivastva (1987). *Appreciative inquiry in organizational life.* Research in Organizational Change and Development, 1, 129-169.

Cowan, G, D. Pines & D. Meltzer. (1994). *Complexity: Metaphors, Models, and Reality.* New York: Addison-Wesley Publishing Company.

Deming, E. (1986). *Out of the Crisis.* Boston: Massachusetts Institute of Technology.

Devore, J. & R. Peck. (1994). *Introductory Statistics Second Edition.* St. Paul, Minnesota: West Publishing Company.

Dooley, K. (1996). A nominal definition of complex adaptive systems. *The Chaos Network*, v8, n1, 2-3.

Dooley, K. (1997). A complex adaptive systems model of organizational change. *Nonlinear Dynamics, Psychology, and Life Sciences*, v1, n1, 69-97.

Eoyang, G. (1997). *Coping with Chaos: Seven Simple Tools.* Cheyenne, Wyoming: Lagumo Publishing.

Fetterman, D., S. Kaftarian, & A. Wandersman (Eds.). (1996). *Empowerment evaluation: Knowledge and tools for self-assessment and accountability.* Thousand Oaks, CA: Sage.

Fink, A. & J. Kosecoff (1996). *How to conduct surveys: A step-by-step guide.* Newbury Park, CA: Sage. Gell-Mann, M. (1994). The Quark and the Jaguar. New York: W. H. Freeman and Company.

Gleick, J. (1988). *Chaos: Making a New Science.* New York: Viking Penguin.

Goldstein, J. (1994). *The Unshackled Organization.* New York: Productivity Press.

Gray, P. (ed.) (1994). *Decision Support and Executive Information Systems.* Englewood Cliffs, New Jersey: Prentice Hall.

Guastello, S. (1995). Chaos, Catastrophe, and Human Affairs: Applications of Nonlinear Dynamics to Work, Organizations, and Social Evolution. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Guba, E. & Y. Lincoln. (1989). *Fourth Generation Evaluation.* Newbury Park, California: SAGE Publications.

Guess, D. & W. Sailor. (1993). Chaos theory and the study of human behavior: Implications for special education and developmental disabilities. *The Journal of Special Education*, v 27, n1, 16-34.

Haynes, S. (1995). Introduction to the special section on chaos theory and psychological assessment. *Psychological Assessment*, v7, n1, 3-4.

Haynes, S., D. Blaine, & K. Meyer. (1995). Dynamical models for psychological assessment: Phase space functions. Psychological Assessment, v7, n1, 17-24.

Heiby, E. (1995a). Assessment of behavioral chaos with a focus on transitions in depression. *Psychological Assessment*, v7, n1, 10-16.

Heiby, E. (1995b). Chaos theory, nonlinear dynamical models, and psychological assessment. *Psychological Assessment*, v7, n1, 5-9.

Kantz, H. & T. Schreiber. (1997). *Nonlinear Time Series Analysis.* New York: Cambridge University Press.

Kaplan, D & L. Glass. (1995). *Understanding Nonlinear Dynamics.* New York: Springer-Verrlag.

Kauffman, S. (1995). *At Home in the Universe.* New York: Oxford University Press.

Kiel, L.D. (1994). Managing Chaos and Complexity in Government: A New Paradigm for Managing Change, Innovation, and Organizational Renewal. San Francisco: Jossey-Bass.

Kellert, S. (1993). *In the Wake of Chaos.* Chicago: The University of Chicago Press.

Kelly, K. (1994). *Out of Control.* New York: Addison-Wesley Publishing Company.

Kerlinger. F. (1985). *Foundations of Behavioral Research.* Third Edition. New York: Holt, Rinehart and Winston, Inc.

Kibel, B. (1996). *Evaluation using Results Mapping.* New Designs for Youth Development, (Winter), pp. 9- 15.

Kontopoulos, K. (1993). *The Logics of Social Structure.* New York: Cambridge University Press.

Lewin, R. (1992). *Complexity: Life at the Edge of Chaos.* New York: Macmillan.

Lorenz, E. (1993). *The Essence of Chaos.* Seattle, Washington: University of Washington Press.

Mainzer, K. (1994). Thinking in Complexity: The Complex Dynamics of Matter, Mind, and Mankind. New York: Springer-Verlag.

March, J. (1994). A Primer on Decision Making: How Decisions Happen. New York: The Free Press.

Morgan, G. (1997). *Images of Organization Second Edition.* Thousand Oaks, California: SAGE Publications, Inc.

Morrow, K. (1990). Evaluating Communicative Tests. Annual Meeting of the Regional Language Centre Seminar. Singapore.

Olson, D. and J. Courtney. (1992). *Decision Support Models and Expert Systems.* New York: Macmillan Publishing Company.

Pandit, S. & Wu, S. (1983). *Time Series and Systems Analysis with Applications.* New York: John Wiley and Sons.

Patton, M. (1990). *Qualitative Evaluation and Research Methods Second Edition.* Newbury Park, California: SAGE Publications, Inc.

Patton, M. (1997). *Utilization-Focused Evaluation: The new century text.* Edition 3. Newbury Park, CA: Sage.

Paulson, F. & P. Paulson. (1991). The ins and outs of using portfolios to assess performance. Revised. Joint Annual Meeting of the National Council of Measurement in Education and the National Association of Test Directors. Chicago.

Prigogine, I. and Stengers, I. (1988). *Order Out of Chaos.* New York: Bantam New Age Books.

Richey, R. (1994). Design 2000: Theory-based design models of the future. Proceedings of Selected Research and Development Presentations at the 1994 National Convention of the Association for Educational Communications and Technology. Nashville, Tennessee.

Stacey, R. (1993). *Managing the Unknowable.* New York: Jossey-Bass.

Stacey, R. (1996). *Creativity and Complexity in Organizations.* San Francisco: Berrett-Koehler Publishers.

Strauss, A. & J. Corbin (1990). Basics of Qualitative Research: Grounded Theory Procedures and Techniques. Newbury Park, CA: Sage.

Stringer, E. (1996). Action research: A Handbook for Practitioners. Thousand Oaks, CA: Sage.

Vaill, P. (1996). Learning as a Way of Being: Strategies for Survival in a World of Permanent White Water. New York: Jossey-Bass.

Van de Ven, A. (1993). The emergence of an industrial infrastructure for technological innovation. *Journal of Comparative Economics*, v17, 338-365.

Waldrop, M. M. (1992). Complexity: The Emerging Science at the Edge of Order and Chaos. New York: Simon & Schuster.

Weick, K. (1979). *The Social Psychology of Organizing.* New York: Random House.

West, B & W. Deering. (1995). *The Lure of Modern Science: Fractal Thinking.* River Edge, New Jersey: World Scientific Publishing Company.

Wheatley, M. (1992). *Leadership and the New Science.* San Francisco: Berrett-Koehler Publishers.

Wheatley, M. (1997). *A Simpler Way.* San Francisco: Berrett-Koehler Publishers.